

# Colloid Science

## *Principles, Methods and Applications*

Edited by

**Terence Cosgrove**

Department of Chemistry  
University of Bristol  
Bristol, UK



**Blackwell**  
Publishing

© 2005 Bristol Colloid Centre

Blackwell Publishing Ltd

Editorial Offices:

Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK

Tel: +44 (0)1865 776868

Blackwell Publishing Professional, 2121 State Avenue, Ames, Iowa 50014-8300, USA

Tel: +1 515 292 0140

Blackwell Publishing Asia Pty Ltd, 550 Swanston Street, Carlton, Victoria 3053, Australia

Tel: +61 (0)3 8359 1011

The right of the Author to be identified as the Author of this Work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

First published 2005 by Blackwell Publishing Ltd

Library of Congress Cataloging-in-Publication Data is available

ISBN-10: 1-4051-2673-6

ISBN-13: 978-14051-2673-1

A catalogue record for this title is available from the British Library

Set in 10/12pt Minion

by Techbooks Electronic Services, India

Printed and bound in India

by Replika Press Pvt. Ltd, India

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on Blackwell Publishing, visit our website:

[www.blackwellpublishing.com](http://www.blackwellpublishing.com)

# Contributors

<b>Paul Bartlett</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Terence Cosgrove</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Sean Davis</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>John Eastman</b>	IMPACT Faraday Ltd, Science & Technology Centre, Whiteknights Road, Reading, RG6 6BZ, UK
<b>Julian Eastoe</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Roy Hughes</b>	Bristol Colloid Centre, University Gate East, Park Row, Bristol, BS1 5UB, UK
<b>Jonathan Reid</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Paul Reynolds</b>	Bristol Colloid Centre, University Gate East, Park Row, Bristol, BS1 5UB, UK
<b>Robert Richardson</b>	Department of Physics, HH Wills Physics Laboratory, University of Bristol, Bristol, BS8 1TL, UK
<b>Jason Riley</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Jeroen van Duijneveldt</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
<b>Brian Vincent</b>	School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK

# Preface

This book has grown out of the *Spring School in Colloid Science*, a one-week course which has been run in the School of Chemistry at Bristol University, during the Easter vacation, every year since 1972. Indeed, this is ‘the book of the course’, since its contents form the basis of the material that is delivered in the lectures. Like the course, this book is primarily intended as a basic introduction to colloid and interface science for those with a first degree in chemistry or physics, or some closely related discipline (e.g. pharmacy, biochemistry), who are working in an industrial R&D laboratory, or indeed a government or university laboratory, and who need to learn the basics of this subject for their research. Sadly, this is a subject which often receives scant coverage still in many undergraduate courses, despite its great relevance in a wide range of chemical technologies, including pharmaceuticals, agrochemicals and food, personal care and household products, surface coatings, oil and mineral recovery and processing. In addition to these more traditional industries, colloid and interface science underpins many of the new, so-called nanotechnologies, such as sensors, IT chips, displays, photonics and micro-reactors.

The authors of the various chapters in this book are all current members of the Colloid Group at Bristol University and are recognised experts in their own areas of the subject. The book begins with an introductory chapter on what a colloid is, how colloids are made and this theme is amplified in Chapter 2 by a detailed discussion of the origin of surface charge that is key to the stability of many aqueous colloidal dispersions. In Chapter 3 the issue of colloid stability is addressed taking into account both the electrostatic and dispersion interactions. In Chapters 4 and 5 two important examples of colloidal systems, surfactants and microemulsions, are discussed in depth. Polymers are also a key feature in colloid science in that they can form aggregates themselves but are widely used in dispersions to promote or destroy stability and in Chapters 6, 7 and 8 their role is treated in detail. Chapter 9 focuses on another aspect of colloid science i.e. the wetting of interfaces and in Chapter 10 some of these ideas are extended to study aerosols which are a further example of a colloidal system. The final three chapters focus on experimental techniques, including rheology in Chapter 11, probably the most basic tool of the colloid scientist, and scattering and imaging methodology in Chapters 12 and 13 and 14 for detailed characterisations of colloidal dispersions.

Although this book is based on the course at Bristol, it should provide a balanced account of the subject from a practical point of view for anyone requiring an introduction to the important world of colloid and interface science.

We would like to acknowledge the following people who helped in the production of this book. Yan Zhang who redrew many of the diagrams, Pam Byrt who helped with the preparation of the manuscripts, Edward Elsey who helped with the proof reading and all the staff of the Bristol Colloid Centre for their support of the project. The authors would especially like to thank their long suffering other halves without whose help and support this book would never have been finished.

# Introduction

Colloid science has a rich history at Bristol, dating back to the early part of the twentieth century, when J.W. McBain was appointed to the chemistry staff in 1907. He was subsequently appointed to the first Leverhulme Chair in Physical Chemistry in 1919, largely through his pioneering work on molecular association in soap solutions (what Hartley later termed 'micelles'), an area in which Lord Leverhulme clearly had a strong interest. McBain left for the United States (to Stanford University) in 1926, whereupon W.E. Garner, whose forté was heterogeneous catalysis, took up the Leverhulme Chair. He was succeeded in turn by D.H. Everett (1954, whose speciality was interfacial and colloidal thermodynamics), R.H. Ottewill (1982, with a broad interest in many aspects of colloid science) and by B. Vincent in 1992.

The first Spring School was conceived by Ron Ottewill, who came to Bristol from Cambridge in 1964, at the instigation of Douglas Everett, primarily to set up the M.Sc. Course by Advanced Teaching and Research in Colloid and Interface Science. This highly successful course ran for some 30 odd years, and only ceased when the introduction of 4-year undergraduate masters degrees in science in the United Kingdom made 1-year post graduate courses less relevant. Many leading industrial scientists and academics in the colloid and interface field have passed through Bristol, as Ph.D. students, members of the M.Sc. course or as participants (around 1000 to date) of the Spring School. Besides Ron Ottewill, other teachers on the first course in 1972 were Aitken Couper, Jim Goodwin, Dudley Thompson and Brian Vincent. The course was initially managed by the Department of Extra-Mural Studies (as it was then called) of the University, and a number of persons from that department gave a great deal of help and support in the early days. Special mention must be made of David Wilde and Sue Pringle in this regard. Since the mid-1990s the Bristol Colloid Centre (BCC) has taken over the management of the course, and it has become one of the range of courses that the BCC offers annually. For example, besides the Spring School, the BCC also offers more basic courses at the technician level and for those without a first degree in a scientific discipline. The BCC was set up in 1994 by Brian Vincent and Jim Goodwin, with Terry Cosgrove subsequently succeeding Jim as Deputy Director after he retired in 1996. Cheryl Flynn was the first staff member to be appointed followed by Paul Reynolds (as manager). The staff now number 11 and the management has been taken on by Roy Hughes with Paul being responsible for new business ventures.

The BCCs purpose is to give research and training support to the large range of industries, in the United Kingdom and abroad, featuring colloid and interfacial science and technology

in their products or processing. The BCC staff and the academic colloid group work very closely together. From the BCC sprung the DTI/EPSRC funded IMPACT (Innovative Manufacturing and Processing using Applied Colloid Technology) Faraday Partnership in 2001, and on the back of IMPACT came the ACORN (Applied Colloid Research Network) DTI Link programme. The BCC has worked closely alongside IMPACT in recently producing its own *on-line* courses in colloid science and technology. The Spring School Course is complementary to these. In particular, together with the lectures, there is a strong emphasis in the Spring School on exposure of the participants to a range of basic experimental techniques. It is out of these beginnings that this text has emerged.

Although the Spring School course, from the start, has been aimed primarily at industrial scientists, the objective has always been to teach the fundamentals of the subject, with many illustrations of the basic principles taken from real applications in industry. Of course, these basics have not changed over the years, but new ideas do strongly emerge and develop, and the course has been updated to accommodate these.

# Contents

<i>Contributors</i>	xii
<i>Preface</i>	xiii
<i>Introduction</i>	xv
1 Introduction to Colloidal Dispersions <i>Brian Vincent</i>	1
1.1 Introduction	1
1.2 Some basic definitions	3
1.2.1 Polydispersity	3
1.2.2 Particle concentration	4
1.2.3 Average particle separation	5
1.2.4 Interfacial area	5
1.3 Interfacial structure	5
1.4 Preparation of colloidal dispersions	6
1.4.1 Comminution	6
1.4.2 Nucleation and growth	8
1.5 Properties of dilute dispersions	9
1.6 Properties of concentrated dispersions	11
1.7 Control of colloid stability	12
References	13
General textbooks in colloid and surface science	13
2 Charge in Colloidal Systems <i>Jason Riley</i>	14
2.1 Introduction	14
2.2 The origin of surface charge	14
2.2.1 Ionisation of surface groups	15
2.2.2 Ion adsorption	15
2.2.3 Dissolution of ionic solids	16
2.2.4 Isomorphous substitution	16
2.2.5 Potential determining ions	16



---

2.3	The distribution of inert ions at a charged interface	16
2.3.1	The mercury/electrolyte interface	17
2.3.2	The Helmholtz model	19
2.3.3	Gouy–Chapman theory	22
2.3.4	The Stern modification	25
2.3.5	Specific adsorption	28
2.3.6	Inter-particle forces	28
2.4	Electrokinetic properties	29
2.4.1	Electrolyte flow	30
2.4.2	Streaming potential measurements	30
2.4.3	Electro-osmosis	31
2.4.4	Electrophoresis	32
2.4.5	Electroacoustic technique	35
	References	35
3	Colloid Stability <i>John Eastman</i>	36
3.1	Introduction	36
3.2	The colloidal pair potential	36
3.2.1	Attractive forces	37
3.2.2	Electrostatic repulsion	38
3.2.3	Effect of particle concentration	40
3.2.4	Total potential	40
3.3	Criteria for stability	42
3.3.1	Salt concentration	42
3.3.2	Counter-ion valency	43
3.3.3	$\zeta$ -potential	44
3.3.4	Particle size	45
3.4	Kinetics of coagulation	46
3.4.1	Diffusion limited rapid coagulation	46
3.4.2	Interaction limited coagulation	47
3.4.3	Experimental determination of the ccc	48
3.5	Conclusions	49
	References	49
4	Surfactant Aggregation and Adsorption at Interfaces <i>Julian Eastoe</i>	50
4.1	Characteristic features of surfactants	50
4.2	Classification and applications of surfactants	51
4.2.1	Types of surfactants	51
4.2.2	Surfactant uses and development	53
4.3	Adsorption of surfactants at interfaces	55
4.3.1	Surface tension and surface activity	55
4.3.2	Surface excess and thermodynamics of adsorption	56
4.3.3	Efficiency and effectiveness of surfactant adsorption	60

---

4.4	Surfactant solubility	62
4.4.1	The Krafft temperature	62
4.4.2	The cloud point	63
4.5	Micellisation	63
4.5.1	Thermodynamics of micellisation	63
4.5.2	Factors affecting the CMC	67
4.5.3	Structure of micelles and molecular packing	68
4.6	Liquid crystalline mesophases	70
4.6.1	Definition	70
4.6.2	Structures	72
4.6.3	Phase diagrams	73
	References	74
5	Microemulsions <i>Julian Eastoe</i>	77
5.1	Microemulsions: definition and history	77
5.2	Theory of formation and stability	78
5.2.1	Interfacial tension in microemulsions	78
5.2.2	Kinetic instability	79
5.3	Physicochemical properties	81
5.3.1	Predicting microemulsion type	81
5.3.2	Surfactant film properties	85
5.3.3	Microemulsion phase behaviour	90
	References	96
6	Polymers and Polymer Solutions <i>Terence Cosgrove</i>	98
6.1	Introduction	98
6.2	Polymerisation	98
6.2.1	Condensation	99
6.2.2	Free radical	99
6.2.3	Ionic methods	99
6.3	Copolymers	100
6.4	Polymer physical properties	100
6.5	Polymer Uses	102
6.6	Theoretical models of polymer structure	102
6.6.1	Radius of gyration	103
6.6.2	Worm-like chains	104
6.6.3	Radius of gyration in ideal solution	104
6.6.4	Excluded volume	104
6.6.5	Scaling theory: blobs	105
6.6.6	Polyelectrolytes	106
6.7	Measuring polymer molecular weight	106
6.8	Polymer solutions	108
	References	112

---

7	Polymers at Interfaces	<i>Terence Cosgrove</i>	113
7.1	Introduction		113
7.1.1	Steric stability		113
7.1.2	The size and shape of polymers in solution		114
7.1.3	Adsorption of small molecules		116
7.2	Adsorption of polymers		117
7.2.1	Configurational entropy		117
7.2.2	The Flory surface parameter $\chi_s$		117
7.3	Models and simulations for terminally attached chains		118
7.3.1	Atomistic modelling		118
7.3.2	Exact enumeration: terminally attached chains		119
7.3.3	Approximate methods: terminally attached chains		121
7.3.4	Scaling models for terminally attached chains		122
7.3.5	Physically adsorbed chains: Scheutjens and Fleer theory		123
7.3.6	Scaling theory for physical adsorption		126
7.4	Experimental aspects		128
7.4.1	Volume fraction profiles		128
7.4.2	Adsorption isotherms		128
7.4.3	The bound fraction		131
7.4.4	The layer thickness		133
7.5	Copolymers		136
7.5.1	Liquid/liquid interfaces		139
7.6	Polymer brushes		140
7.7	Conclusions		141
	References		141
8	Effect of Polymers on Colloid Stability	<i>Jeroen van Duijneveldt</i>	143
8.1	Introduction		143
8.1.1	Colloid stability		143
8.1.2	Limitations of charge stabilisation		143
8.1.3	Effect of polymers on interactions		144
8.2	Particle interaction potential		144
8.2.1	Measuring surface forces		144
8.3	Steric stabilisation		145
8.3.1	Theory		145
8.3.2	Steric stabiliser design		148
8.3.3	Marginal solvents		149
8.4	Depletion interactions		151
8.5	Bridging interactions		155
8.6	Conclusion		157
	References		157
9	Wetting of Surfaces	<i>Paul Reynolds</i>	159
9.1	Introduction		159
9.2	Surfaces and definitions		159

---

9.3	Surface tension	160
9.4	Surface energy	161
9.5	Contact angles	161
9.6	Wetting	162
9.7	Liquid spreading and spreading coefficients	164
9.8	Cohesion and adhesion	165
9.9	Two liquids on a surface	166
9.10	Detergency	168
9.11	Spreading of a liquid on a liquid	170
9.12	Characterisation of a solid surface	172
9.13	Polar and dispersive components	172
9.14	Polar materials	173
9.15	Wettability envelopes	174
9.16	Measurement methods	176
9.17	Conclusions	178
	References	178
10	Aerosols <i>Jonathan Reid</i>	180
10.1	Introduction	180
10.2	Generating and sampling aerosol	184
10.2.1	Generating aerosol	185
10.2.2	Sampling aerosol	186
10.3	Determining particle concentration and size	188
10.3.1	Determining number concentration	189
10.3.2	Determining mass concentration	189
10.3.3	Determining particle size	190
10.4	Determining particle composition	195
10.4.1	Sampling and off-line analysis	195
10.4.2	Real-time analysis	196
10.4.3	Single particle analysis	197
10.5	Concluding remarks	199
	References	200
11	Practical Rheology <i>Roy Hughes</i>	201
11.1	Introduction	201
11.2	Making measurements	201
11.2.1	Definitions	201
11.2.2	Designing an experiment	203
11.2.3	Geometries	206
11.2.4	Viscometry	207
11.2.5	Shear thinning and thickening behaviour	210
11.3	Rheometry and viscoelasticity	212
11.3.1	Viscoelasticity and Deborah number	212
11.3.2	Oscillation and linearity	212
11.3.3	Creep compliance	214

---

11.3.4	Liquid and solid behaviour	214
11.3.5	Sedimentation and storage stability	216
11.4	Examples of soft materials	219
11.4.1	Simple particles and polymers	220
11.4.2	Networks and functionalisation	221
11.4.3	Polymeric additives	223
11.4.4	Particle additives	225
11.5	Summary	227
	References	227
12	Scattering and Reflection Techniques <i>Robert Richardson</i>	228
12.1	Introduction	228
12.2	The principle of a scattering experiment	229
12.3	Radiation for scattering experiments	230
12.4	Light scattering	231
12.5	Dynamic light scattering	233
12.6	Small angle scattering	234
12.7	Sources of radiation	235
12.8	Small angle scattering apparatus	235
12.9	Scattering and absorption by atoms	237
12.10	Scattering length density	237
12.11	Small angle scattering from a dispersion	239
12.12	Form factor for spherical particle	240
12.13	Determining particle size from SANS and SAXS	240
12.14	Guinier plots to determine radius of gyration	241
12.15	Determination of particle shape	242
12.16	Polydispersity	242
12.17	Determination of particle size distribution	242
12.18	Alignment of anisotropic particles	244
12.19	Concentrated dispersions	245
12.20	Contrast variation using SANS	246
12.21	High Q limit: Porod law	247
12.22	Introduction to X-ray and neutron reflection	250
12.23	Reflection experiment	250
12.24	A simple example of a reflection measurement	251
12.25	Conclusion	252
	References	253
13	Optical Manipulation <i>Paul Bartlett</i>	255
13.1	Introduction	255
13.2	Manipulating matter with light	255
13.3	Force generation in optical tweezers	258
13.4	Nanofabrication	260

---

13.5	Single particle dynamics	261
13.5.1	Measuring nanometre displacements	261
13.5.2	Brownian fluctuations in an optical trap	262
13.5.3	Dynamical complexity in colloidal gels	263
13.6	Conclusions	265
	References	265
14	Electron Microscopy <i>Sean Davis</i>	266
14.1	General features of (electron) optical imaging systems	266
14.2	Conventional TEM	268
14.2.1	Background	268
14.2.2	Practical aspects	268
14.2.3	Polymer latex particles	269
14.2.4	Core/shell particles	270
14.2.5	Internal structure	272
14.3	Conventional SEM	275
14.3.1	Background	275
14.3.2	Types of signal	275
14.3.3	Practical aspects	275
14.4	Summary	281
	References	282
	<i>Index</i>	283

## Chapter 1

# Introduction to Colloidal Dispersions

*Brian Vincent*

### 1.1 Introduction

This chapter is intended to serve as a basic introduction to colloid science for those with little or no experience of the subject. Some basic concepts and definitions will be discussed, which will lay the foundations for the chapters which follow. References to some other modern textbooks are given in Section 1.8

The simplest definition of a colloidal dispersion is that it is a two-phase system, where one phase (the dispersed phase  $\beta$ ) is dispersed in the second phase (the continuous phase  $\alpha$ ), as illustrated in Figure 1.1.

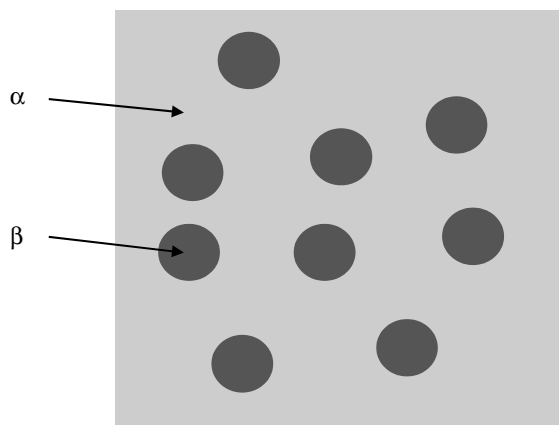
The dispersed phase particles, droplets or bubbles may be spheres (diameter  $d$ ), as shown in Figure 1.1. On the other hand, solid particles could also be cubes, plates, rods, etc. At least one dimension of the particles (i.e.  $d$  for spheres) is usually in the range 1 nm–10  $\mu$ m. In 1917, Wolfgang Ostwald [1] referred to ‘the world of neglected dimensions’ to describe those material objects lying between the molecular and the macroscopic (bulk).

Some common, everyday examples of colloidal dispersions are given in Table 1.1.

In many practical cases, the system may be rather more complex. There may be more than one type of dispersed phase present, and any of the phases (dispersed or continuous) may be multi-component (for example, in an aqueous phase, there may be electrolytes, surfactants, polymers and other molecular species present). Ice cream is an interesting example of a complex colloidal dispersion in that all three types of dispersed phases are present: solid particles (ice), liquid droplets (fat globules) and gas bubbles (air).

Colloid science underpins colloid technology, that is, the formulation of colloidal-based products and processes. Some common examples are listed in Table 1.2.

We have so far discussed colloidal systems in terms of dispersions of one (or more) phases in a continuous phase. Clearly in such systems the role of the (various) *interfaces* between the different phases becomes an important consideration. However, there is another class of colloidal system that should be introduced here, namely association colloids – see Section 1.4. As the name suggests, these are solutions in which molecular association occurs to produce entities in the nanometre size range. A typical example would be surfactant micelles, which are formed by the association of surfactant molecules, above a certain concentration in solution, referred to as the ‘critical micelle concentration’. These are *equilibrium* structures (unlike most multi-phase colloids), having a defined size and shape, under given conditions. A more modern definition would be ‘supramolecular solutions’.



**Figure 1.1** Schematic representation of a colloidal dispersion.

One might also consider a third class of systems as being in the ‘colloidal’ domain; these are solutions of macromolecules, since homopolymer chains in solution typically have dimensions (e.g. radii of gyration) in the nanometre range. In fact, in the early days of the last century, polymers were regarded as colloidal dispersions, until their high molecular weight nature was proposed by Staudinger (1920), for which he received the Nobel Prize in Chemistry in 1953. The outstanding series of experiments by Carothers in 1931 [2] were to confirm this macromolecular viewpoint. The worlds of colloid science and polymer science then separated, although one could argue that more modern objects like polymer dendrimers and block copolymer micelles perhaps span both worlds.

One might also ask the question as to where colloid science and colloid technology stand in relation to ‘nanoscience’ and ‘nanotechnology’. In terms of the *science*, they are both firmly rooted in a common, underpinning science, namely that of surfaces and interfaces. This subject is essentially concerned with how molecules, and assemblies of molecules, behave (both structurally and dynamically) on the nanoscale, at or near interfaces, as opposed to in bulk. Where differences may arise is in the *technology*, i.e. the applications; colloid technology is implicit in all those products and processes listed in Table 1.2, and many similar ones. Although there is some confusion and debate as to what actually constitutes

**Table 1.1** Main types of colloidal dispersions

Medium( $\alpha$ ) $\Rightarrow$	Gas	Liquid	Solid
Dispersed phase ( $\beta$ ) $\Downarrow$			
Gas (Bubbles)	****	Foam Beer head	Solid foam Polystyrene cup
Liquid (Droplets)	Liquid aerosol Mist	Liquid emulsion Mayonnaise	Solid emulsion Butter
Solid (Particles)	Solid aerosol Smoke	Sol Ink	Solid sol Stained glass



**Table 1.2** Colloids in everyday life

Products	Processes
<ul style="list-style-type: none"> <li>• Surface coatings (Paints, video tapes, photographic films)</li> <li>• Cosmetics and personal care (Creams, toothpaste, hair shampoo)</li> <li>• Household products (Liquid detergents, polishes, fabric conditioners)</li> <li>• Agrochemicals (Pesticides, insecticides, fungicides)</li> <li>• Pharmaceuticals (Drug delivery systems, aerosol sprays)</li> <li>• Foodstuffs (Butter, chocolate ice cream, mayonnaise)</li> <li>• Pigmented plastics</li> <li>• Fire-fighting foams</li> </ul> <p>Also natural systems, such as:</p> <ul style="list-style-type: none"> <li>• Biological cells</li> <li>• Mists and fogs</li> </ul>	<ul style="list-style-type: none"> <li>• Clarification of liquids (Water, wine, beer)</li> <li>• Mineral processing (Flotation, selective flocculation)</li> <li>• Detergency (‘Soil’ detachment, solubilisation)</li> <li>• Oil recovery (Drilling fluids, oil slick dispersal)</li> <li>• Engine and lube oils (Dispersion of carbon particles)</li> <li>• Silting of river estuaries</li> <li>• Ceramic processing (‘Sol <math>\Rightarrow</math> gel’ processing)</li> <li>• Road surfacing (Bitumen emulsions)</li> </ul>

nanotechnology, it seems to be more concerned with *miniaturised devices*, such as memory chips, sensors, opto/electrical devices, valves, pumps, motors, etc. However, with regard to the commonality of the *basic science* underpinning colloid technology and nanotechnology, the basic physics of say using titanium dioxide particles in sunscreen lotions (a traditional colloidal application) is very similar to that for their use in miniature opto/electrical devices for converting sunlight into electricity!

## 1.2 Some Basic Definitions

### 1.2.1 Polydispersity

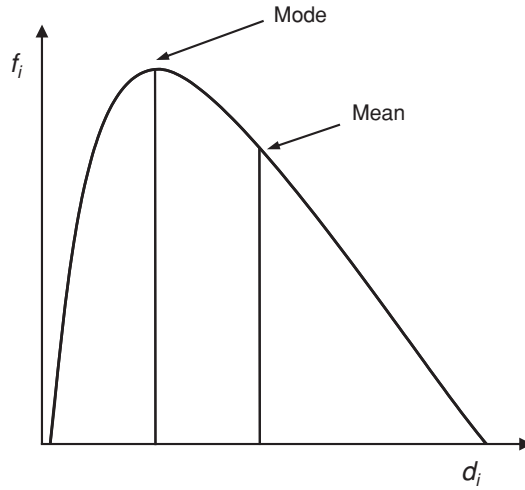
As with macromolecules, all colloidal systems are polydisperse with regard to their size, to a greater or lesser extent. Pure science tends to favour systems which are as ‘monodisperse’ as possible, as model systems, whereas in practical applications a large polydispersity can sometimes bring beneficial properties to a product, for example, preferable rheological properties.

With respect to Figure 1.2 which shows a typical particle size distribution,

$$f_i = \frac{N_i}{N_{\text{tot}}} \quad (1.1)$$

mean:

$$\langle d \rangle = \sum_i f_i d_i \quad (1.2)$$



**Figure 1.2** Typical particle size distribution.

standard deviation:

$$\sigma = \left[ \sum_i f_i (d_i - \langle d \rangle)^2 \right]^{\frac{1}{2}} \quad (1.3)$$

variance:

$$v = \frac{100\sigma}{\langle d \rangle} \quad (1.4)$$

where  $N_i$  is the number of particles of diameter  $d_i$  and  $N_{\text{tot}}$  is the total number of particles.

### 1.2.2 Particle concentration

Several terms may be used to define the particle concentration in a dispersion. For  $N$  particles (particle volume  $v_p$ ) in a total volume of dispersion  $V$ , number density:

$$n = \frac{N}{V} \quad (1.5)$$

volume fraction:

$$\phi = nv_p \quad (1.6)$$

weight fraction:

$$w = \frac{\phi \rho_p}{\rho_s} \quad (1.7)$$

where  $\rho_p$  and  $\rho_s$  are the densities of the particle and the solvent, respectively.

Two particular values of  $\phi$  should be noted,  $\phi_{\text{hcp}} (=0.74)$  and  $\phi_{\text{rcp}} (=0.64)$ , corresponding to hexagonal close-packing and random close-packing of the particles, respectively.

### 1.2.3 Average particle separation

The average (minimum) separation distance ( $h$ ) between two spheres is given by

$$\langle h \rangle = \frac{d}{2} \left[ \left( \frac{\phi_{\text{rcp}}}{\phi} \right)^{\frac{1}{3}} - 1 \right] \quad (1.8)$$

Note that  $\langle h \rangle = d$ , when  $\phi = 2.4\%$ .

### 1.2.4 Interfacial area

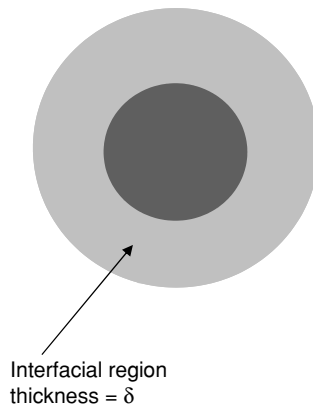
The total area ( $A$ ) of the interface in a colloidal dispersion is given by

$$A = N_{\text{tot}} \pi d^2 = \frac{6V\phi}{d} \quad (1.9)$$

By way of example, for  $V = 1 \ell$ ,  $d = 100 \text{ nm}$  and  $\phi = 0.1$ ,  $A = 6000 \text{ m}^2$ , which is the approximate area of a soccer pitch. This indicates just how important interfaces are in colloidal dispersions, and why interfacial science underpins colloid science.

## 1.3 Interfacial Structure

The structure of the liquid in the interfacial region (see Figure 1.3) is usually somewhat different from that in the bulk, due to the influence of the particle (droplet or bubble) surface. For example, for an aqueous dispersion of *hydrophilic* particles (e.g. most oxides, such as silica) the water molecules are more ‘ordered’ (H-bonded) near the solid surface



**Figure 1.3** Schematic of the interfacial region around a colloidal particle.

than in the bulk. On the other hand, for *hydrophobic* particles (e.g. polystyrene latex), the water molecules are somewhat depleted (in concentration) near the solid surface.

The extent ('thickness') of the interfacial region ( $\delta$ , see Figure 1.3), depends on the system. For dispersions of neutral particles, in a pure solvent or mixture of solvents,  $\delta$  may be of the order of a few molecular diameters. For adsorbed polymer layers,  $\delta$  depends on the conformation of the polymer chains at the surface, but will be of the order of radius of gyration of the polymer in solution.

In aqueous media in particular, particles (droplets or bubbles) develop an interfacial charge. This can arise in several ways:

- (a) For sparingly soluble ionic particles (e.g. AgI), there will be a (different) equilibrium distribution of the constituent ions (i.e.  $\text{Ag}^+$  and  $\text{I}^-$ ) between the solid particles and the solution phase. This arises because one of the ions will invariably be preferentially solvated in solution, in this case the  $\text{Ag}^+$  ion.
- (b) For polymer particles, end groups on the constituent chains (from initiator fragments) may well become de-protonated (e.g.  $-\text{COOH}$ ) or protonated (e.g.  $-\text{NH}_2$ ) under the appropriate pH conditions.
- (c) There may be specific adsorption of charged species added to the solution phase, e.g. simple ions (such as  $\text{Ca}^{2+}$ ), surfactant ions and polyelectrolyte chains. This route is the normal charging mechanism in non-polar media, when an added organic-soluble electrolyte may be used as the 'charging agent' (e.g. tetra alkyl ammonium phenolates or benzoates).
- (d) For certain ionic solid particles, there may be a net charge associated with the particle, due to isomorphic substitution of one constituent ion by another of different valency, e.g. in certain clays the replacement of an  $\text{Al}^{3+}$  ion by a  $\text{Mg}^{2+}$  ion leads to a net negative charge on the clay platelet faces.

As illustrated schematically in Figure 1.4, there is an associated, diffuse layer of (free) ions of opposite charge, which form an ion 'atmosphere' around each charged particle; these ions are called 'counter-ions'. There is an associated electrostatic potential ( $\psi_0$ ) at the particle surface; for a negatively charged surface, this potential is also negative. The concentration ( $c^+$ ) of the counter-ions (in this case cations) in the diffuse layer decays exponentially from the surface, as illustrated in Figure 1.4. The thickness ( $\delta$ ) of this diffuse layer of counter charge, and what controls it, will be discussed in Section 2.3.3.

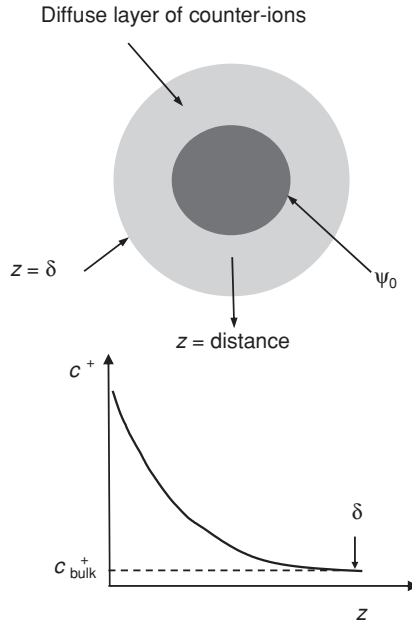
## 1.4 Preparation of Colloidal Dispersions

There are two general methods of producing colloidal dispersions (particles, droplets or bubbles). These may be discussed in terms of Figure 1.5.

### 1.4.1 Comminution

This requires mechanical work to be done on the system. The work done ( $w$ ) on the system is given by

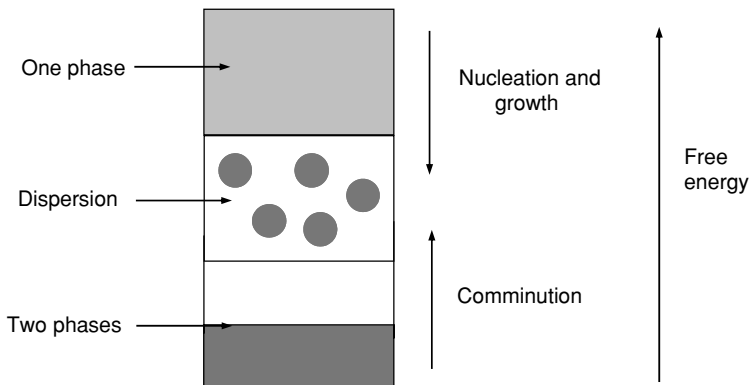
$$w = \Delta F = \gamma \Delta A + q \quad (1.10)$$



**Figure 1.4** The electrical double layer for a charged particle.

where  $\Delta F$  is the associated increase in free energy of the system,  $\gamma \Delta A$  is the interfacial free energy term ( $\gamma$  is the interfacial tension and  $A$  the resulting increase in surface area – see Equation (1.9)) and  $q$  is the heat generated. Examples of comminution processes are:

- (i) when two essentially immiscible phases are subjected to shear (e.g. with a stirrer) or extensional flow (e.g. a fluid passed through a frit into a liquid)
- (ii) when larger particles are subjected to mechanical abrasion (as in a ball mill) to produce smaller particles
- (iii) when a liquid-based dispersion is subjected to a cavitation process, e.g. through application of ultrasonic radiation.



**Figure 1.5** Preparation of dispersions.

### 1.4.2 Nucleation and growth

This occurs when, through a change in some thermodynamic parameter (e.g. temperature  $\Delta T$ , pressure  $\Delta p$ , or concentration  $\Delta c$ , including addition of a new component), an incipient phase change results. Nucleation and growth occur with a decrease in free energy, that is, spontaneously. Some examples of dispersions produced by nucleation and growth are:

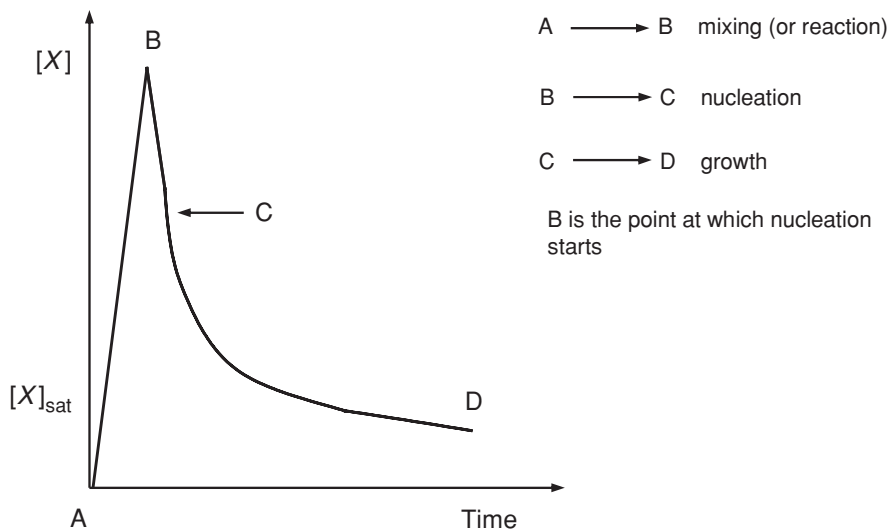
- (i)  $\Delta c$ : gold sols:  $\text{HAuCl}_4 + \text{a reducing agent} \Rightarrow \text{Au}^0 \downarrow$  (c.f. Faraday's gold sol at the RI Museum, London, prepared in 1857) or, polymer latexes: monomer solution + an initiator  $\Rightarrow$  polymer particles
- (ii)  $\Delta p$ : foams, e.g. canned beverages
- (iii)  $\Delta T$ : emulsions (by crossing a phase boundary, from a one-phase region to a two-phase region, in a liquid mixture).

In general, nucleation and growth is a better controlled process than comminution, and better control of the particle size (distribution) may, in general, be achieved. The key to this control is manipulating the state of super-saturation of the nucleating species,  $X$  (see Figure 1.6).

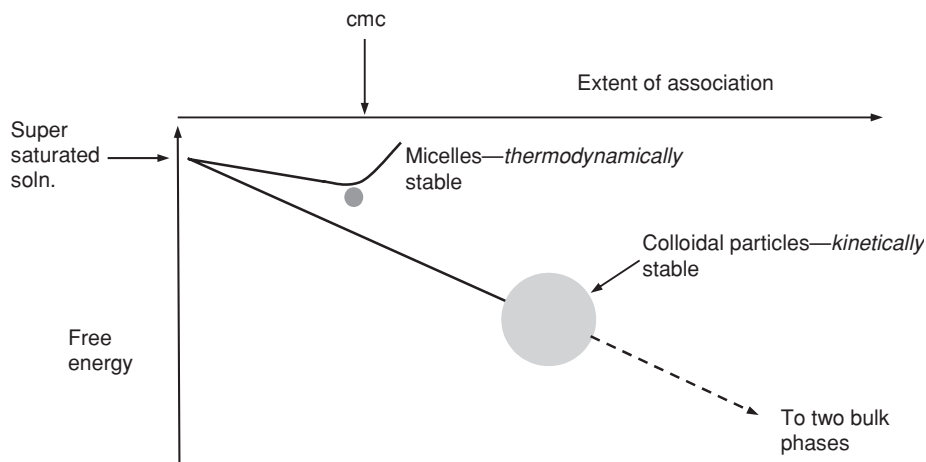
Ideally, for monodisperse particles, all three stages indicated in Figure 1.6 should be distinct, i.e. occur sequentially. This is difficult to achieve in practice.

Note that during the growth stage, in particular, the particles, droplets or bubbles need to be stabilised against coagulation and (for droplets and bubbles) coalescence. This may be achieved by (a) interfacial charge and/or (b) adsorbed polymer or surfactant.

As mentioned earlier, another class of colloidal dispersion is association colloids. Unlike nucleation and growth, discussed above, in this case the 'particles' only grow to a certain size



**Figure 1.6** Nucleation and growth.



**Figure 1.7** Comparison of nucleation and growth and association processes.

(and shape), which is controlled by the amphipathic nature and ‘geometry’ of the molecules concerned.

Figure 1.7 illustrates the difference between these two processes. The smaller particle is e.g. a micelle. As molecular association proceeds, a free energy minimum is reached, which corresponds to the equilibrium micelle size. Hence, micelles are *thermodynamically* stable. The larger particle is a conventional colloidal particle, i.e. a separated dispersed phase. Since the free energy of the system decreases continuously, in principle, this particle could go on growing in size, but at some stage *kinetic* stability is achieved. That means that at a certain size, the charge density reaches a point where the inter-particle repulsive forces effectively exceed the van der Waals attraction between the particles (see Section 1.5d)

## 1.5 Properties of Dilute Dispersions

In a ‘dilute’ dispersion, the particles (droplets or bubbles) are sufficiently isolated that they behave, for most of the time, as individual particles (as do molecules in the gas phase). They are subject to certain forces, however:

- (a) Inertial forces from the solvent molecules, giving rise to Brownian motion. This may be discussed in terms of the diffusion coefficient ( $D$ ) for the particles,

$$D = \frac{kT}{f} \quad (\text{Einstein}) \quad (1.11)$$

$$f = 6\pi\eta a \quad (\text{Stokes – for spherical particles}) \quad (1.12)$$

where  $k$  is the Boltzmann constant,  $T$  is the absolute temperature,  $f$  is the frictional coefficient of the particle in the solvent,  $\eta$  is the solvent viscosity and  $a$  is the particle radius.

- (b) External field forces. In the Earth's gravitational field, the sedimentation velocity ( $v_s$ ) of a particle is given by,

$$v_s = \frac{2\pi \Delta\rho g a^2}{9\eta} \quad (1.13)$$

where  $\Delta\rho$  is the density difference between the particle and the medium, and  $g$  is the gravitational constant. By comparing Equations (1.11) and (1.12) with (1.13), it may be deduced that diffusion will dominate sedimentation for small particle sizes, whereas the reverse is true for larger particles.

In an applied electrical field the particle will undergo electrophoretic motion, where the electrophoretic velocity ( $v_{el}$ ) is given by

$$v_{el} = \frac{2\sigma a E}{3\eta} \quad (1.14)$$

where  $\sigma$  is the surface charge density of the particles and  $E$  is the applied electric field. Note that the electrophoretic mobility ( $u$ ) is given by  $v_{el}/E$ .

- (c) Shear forces. In this case the effect is usually observed as a change in the properties of the whole dispersion, rather than the particles themselves. For example, the shear viscosity of a dispersion depends on the particle volume fraction ( $\phi$ ), for dilute dispersions of spherical particles, as follows:

$$\eta = \eta_0(1 + 2.5\phi) \quad (\text{Einstein}) \quad (1.15)$$

where  $\eta_0$  is the viscosity of the continuous phase.

- (d) Inter-particle forces. There are always long-range attractive forces between the constituent particles in a dispersion; these are due to van der Waals forces and are discussed in more detail in Chapter 3. In addition, there are what might, somewhat loosely, be described as 'structural' forces between the particles. These may best be understood in terms of Figure 1.3, where the structurally different, interfacial region or sheath surrounding all colloidal particles was described. When any two particles approach, there will necessarily be an 'overlap' of their respective interfacial sheaths. Structural perturbations of these interfacial sheaths must then occur (at the very least a displacement of solvent molecules from the 'overlap zone' into the bulk). This overlap will be accompanied by either an *increase* in free energy (i.e. a *repulsive interaction*) or a *decrease* in free energy (i.e. an *attractive interaction*), or possibly both, at different particle separations. The details of this will be discussed in later chapters. However, one such structural interaction is the electrostatic repulsion between (like) charged particles which results from overlap of their electrical double layers. The combination of electrostatic repulsion (electrolyte concentration dependent) and van der Waals inter-particle attraction formed the basis of the classical, so-called Derjaguin–Landau–Verwey–Overbeek (DLVO) theory of the stability of charge-stabilised particulate dispersions, first proposed by Derjaguin and Landau in 1941 [3] in Russia, and independently by Verwey and Overbeek in 1948 in Holland [4]. At low electrolyte concentrations, electrical double-layer repulsion dominates, so the dispersion is kinetically stable. As the electrolyte concentration is gradually increased, the electrical double layer repulsion is diminished (in magnitude and range) and so the inter-particle van der Waals attraction becomes more dominant, and the rate



of coagulation of the particles increases. We shall return to a further, brief discussion of particle aggregation processes in Section 1.7.

- (e) Effects of interfacial tension (which are associated with the imbalance of the intermolecular forces near an interface). These give rise to two principle instabilities in colloidal dispersions:

- (i) coalescence (at least for droplets and bubbles). By reducing the total interfacial area in the dispersion (see Equation (1.9)), the free energy of the dispersion ( $F$ ) is reduced, i.e.

$$\Delta F = \gamma \Delta A \quad (1.16)$$

- (ii) Ostwald ripening [4]. This process is associated with the increase in Laplace pressure ( $\Delta p$ ) inside a spherical particle, droplet or bubble, with increasing curvature, i.e. decreasing radius), i.e.

$$\Delta p = \frac{2\gamma}{a} \quad (1.17)$$

As a result of this pressure difference, the chemical potential of the molecules constituting the particles (droplets or bubbles) decreases with increasing particle size. This means that larger particles tend to grow at the expenses of smaller particles, provided there is some (limited) solubility of these molecules in the continuous phase.

## 1.6 Properties of Concentrated Dispersions

It is, of course, difficult to define a boundary *per se* between dilute and concentrated dispersions. A working hypothesis might be to consider a dispersion to be ‘concentrated’ when, with increasing  $\phi$ , the time between particle collisions decreases to an extent where that time becomes comparable to the time spent together as a collision event – again this is difficult to define, but typically it is of the order of microseconds.

One should bear in mind that, in terms of particle interactions and collisions, it is the total, *effective* radius of the particle that is important, i.e. the core radius *plus* the interfacial sheath thickness (i.e.  $a + \delta$ ), see Figure 1.3. One may, therefore, define an effective particle volume fraction ( $\phi_{\text{eff}}$ ) as follows:

$$\phi_{\text{eff}} = \phi \frac{(a + \delta)^3}{a^3} \quad (1.18)$$

In practice, a dispersion becomes ‘concentrated’ when the effective particle volume fraction is increased to somewhere in the range  $0.01 < \phi_{\text{eff}} < 0.1$  (or 1–10% by volume).

Many of the properties of dispersions change once they become concentrated (and most commercial products (see Table 1.2) are indeed formulated as concentrated dispersions). For example:

- (a) The particles are strongly interacting, through the inter-particle forces, most of the time, rather than occasionally, as in a dilute dispersion. The interactions become multi-body, rather than pairwise, i.e. the interaction between any two particles is modified by the other neighbouring particles. This is a much more difficult problem to deal with than the simple pairwise additivity assumed for dilute dispersions.

- (b) There will be, as a result of (a), strong spatial correlations between the particles. For example, some dispersions may form colloidal crystals, if they are sufficiently monodisperse and the inter-particle interactions are net repulsive.
- (c) There will be strong hydrodynamic coupling between the particles, such that the simple equations given earlier for particle diffusion, sedimentation, electrophoresis and for the dispersion viscosity become invalid. Often complete theories become difficult to formulate at higher  $\phi$  values, so often what is done is to represent a particular hydrodynamic property by an (empirical) virial expansion in  $\phi$ . So, for example, the sedimentation velocity (see Equation (1.13)) becomes

$$v_s = v_{s,0}(1 + a\phi + b\phi^2 + \dots) \quad (1.19)$$

where  $v_{s,0}$  is the sedimentation velocity in the dilute limit, i.e. as  $\phi \Rightarrow 0$ , and  $a$  and  $b$  (etc.) are virial coefficients that are used to fit the experimental data to a polynomial of the appropriate order.

## 1.7 Control of Colloid Stability

Although this subject will be dealt with in much greater depth elsewhere in this book, it is worth making a few preliminary remarks here. Primarily, in many practical cases, one is faced with one of three main objectives in handling colloidal dispersions:

- (a) To achieve long-term stability of that dispersion, as in pharmaceutical, cosmetic and food products, for example. The dispersion may actually be *thermodynamically* stable. This is the case when the free energy of flocculation ( $\Delta F_{\text{floc}}$ ) is *positive*.  $\Delta F_{\text{floc}}$  is given by

$$\Delta F_{\text{floc}} = \Delta U_{\text{floc}} - T\Delta S_{\text{floc}} \quad (1.20)$$

where  $\Delta U_{\text{floc}}$  and  $\Delta S_{\text{floc}}$  are the energy and entropy of flocculation, respectively. They are both negative quantities.  $\Delta U_{\text{floc}}$  is related to the energy per contact-pair ( $u$ ) in a floc, and  $\Delta S_{\text{floc}}$  is related to  $\phi$  (the lower  $\phi$ , the greater the loss in entropy when particles flocculate). So, if  $T\Delta S_{\text{floc}}$  is  $> \Delta U_{\text{floc}}$  in magnitude, then  $\Delta F_{\text{floc}}$  is positive overall and the dispersion will indeed be thermodynamically stable. In general,  $u$  has to be rather small to achieve this condition. In practice, however, long-term *kinetic stability* will usually suffice.

- (b) To achieve rapid, strong aggregation of a dispersion. Examples here would include clarification processes (e.g. of water or beer) or ‘harvesting’ processes (e.g. of minerals or bacterial growth cultures) where rapid aggregation increases the settling rate of the particles.
- (c) The most challenging objective for colloid scientists, however, is to be able to control the extent and energy of flocculation, i.e. to achieve weak, reversible aggregation. This is important in controlling:
  - (i) the rheological properties of a dispersion (e.g. to make the system shear-thinning and/or have a yield stress)
  - (ii) the nature of any sedimented particles (i.e. to make them re-dispersible on simple shaking or stirring).

All this is primarily achieved by controlling (for a given value of  $\phi$ ) the value of the pair contact energy ( $u$ ). If  $u$  is less than say  $5kT$ , then weak, reversible flocculation may be achieved. This demands a much deeper understanding of the various types of inter-particle interactions that can occur, in particular when *polymers* are added to dispersions. However, this is beyond the scope (and intended objectives) of this introductory chapter.

## References

1. Ostwald, W.O. (1917) *An Introduction to Theoretical and Applied Colloid Chemistry*. Wiley, New York.
2. Carothers, W.H. (1931) *Ber. Deutsch. Chem. Ges. B*, 53, 1073–1085.
3. Derjaguin, B.V. and Landau, L.D. (1944) *Acta Physiochim. URSS*, 14, 633–662.
4. Verwey, E.J.W. and Overbeek, J.Th.G. (1948) *Theory of the Stability of Lyophobic Colloids*. Elsevier, Amsterdam, (2000) Dover reprint.

## General Textbooks in Colloid and Surface Science

- Adamson, A.W. and Gast, A.P. (1997) *Physical Chemistry of Surfaces*. Wiley, New York.
- Evans, D.F. and Wennerstrom, H. (1998) *The Colloidal Domain: Where Physics, Chemistry and Biology Meet*. Wiley, New York.
- Goodwin, J.W. (2004) *Colloids and Interfaces, with Surfactants and Polymers: An Introduction*. Wiley, New York.
- Hiemenz, P.C. and Rajagopalan, R. (1997) *Principles of Colloid and Surface Chemistry*. Dekker, New York.
- Holmberg, K. (ed.) (2001) *Applied Surface and Colloid Chemistry* (vols 1 and 2). Wiley, New York.
- Hunter, R.J. (1993) *Introduction to Modern Colloid Science*. Oxford University Press, Oxford.
- Lyklema, J. (1991–2005) *Fundamentals of Interface and Colloid Science* (vols 1–5). Academic Press, New York.
- Morrison, I.D. and Ross, S. (2002) *Colloidal Dispersions: Suspensions, Emulsions and Foams*. Wiley, New York.
- Norde, W. (2003) *Colloids and Interfaces in Life Sciences*. Dekker, New York.
- Pashley, R.M. (2004) *Applied Colloid and Surface Chemistry*. Wiley, New York.

## Chapter 2

# Charge in Colloidal Systems

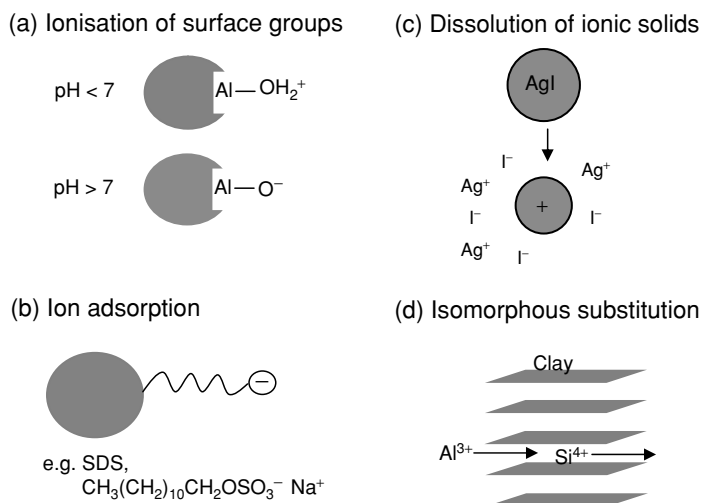
*Jason Riley*

### 2.1 Introduction

Between particles in any suspension there exists, due to induced dipoles, an attractive van der Waals interaction. This attractive force promotes aggregation of particles. To prepare stable colloidal suspensions it is necessary to introduce interactions between particles that oppose the van der Waals attraction. One method of achieving this is to charge the particles, the surface charge resulting in a repulsive inter-particle force [1]. Methods of charging particles can be classified under four broad headings; ionisation of surface groups, ion adsorption, non-symmetric ion dissolution and isomorphous ion substitution. In an electrolyte the ions surround the particles and shield their surface charge. The distribution of counterions in the vicinity of a charged surface may be described using Stern–Gouy–Chapman theory in which the potential at the surface is dropped across two layers, a compact inner layer and a diffuse outer layer. As two particles approach each other their diffuse layers will overlap and the resultant repulsive force may outweigh the attractive van der Waals attraction, thus rendering the suspension stable. The distribution of ions in the diffuse layer is dependent upon the concentration of the electrolyte, the formal charge of the ions, the solvent and the potential at the boundary between the compact inner layer of ions and the diffuse outer layer of ions. The potential at this interface is often equated to the zeta ( $\zeta$ ) potential, that is the potential at the shear plane between the particle and the solvent under flow. The  $\zeta$ -potential is determined in electrokinetic experiments; techniques in which relationships between the current or voltage and the relative flow of the two phases in the suspension are measured. Hence, the scientist wishing to prepare stable colloids must have an appreciation of the origins of surface charge, understand how ions are distributed at the particle/liquid interface and be familiar with methods of measuring the  $\zeta$ -potential.

### 2.2 The Origin of Surface Charge

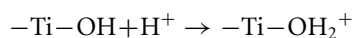
Many methods of preparing colloidal suspensions yield particles that possess charge at their surface. The surface charge may be further modified by altering the environment, for example changing the pH or adding an ionic surfactant. There are four generic mechanisms by which a surface immersed in a liquid may attain a charge; these are summarised in Figure 2.1.



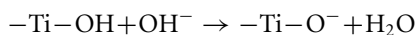
**Figure 2.1** The methods of charging a solid surface immersed in electrolyte.

### 2.2.1 Ionisation of surface groups

Particles that possess suitable chemical functionality can become charged as a result of the ionisation of surface groups. In aqueous solutions pH is commonly used to control the degree and nature of the ionisation. For example metal oxides may become charged as a result of the protonation or deprotonation of surface groups. The pH of the isoelectric point of titania is 5.8, i.e. at a pH of 5.8 the  $\zeta$ -potential of titania is zero. At a pH less than 5.8 the titania is positively charged:



and at a pH greater than 5.8 the oxide is negatively charged:



Similarly proteins may acquire charge as a result of the ionisation of carboxyl and amino groups to  $-\text{COO}^-$  and  $-\text{NH}_3^+$  respectively.

### 2.2.2 Ion adsorption

If the bulk material cannot be ionised, ionic surfactants may be added to generate charge stabilised suspensions. For example, particles of carbon black on which anionic surfactants are adsorbed may be suspended in water; this is the basis of inks and the proprietary mixture of surfactants used is termed a dispersant. Increases in pH may result in protonation of the anionic surfactants of the dispersant and lead to instability of the inks.

### 2.2.3 Dissolution of ionic solids

Silver halide sols underpin photographic film technology and have played an important role in the development of our understanding of charged colloids. Silver halides are sparingly soluble salts, for example, the solubility product of silver iodide ( $K_{SP} = a_{\text{Ag}^+} a_{\text{I}^-}$ ) in pure water is  $8.5 \times 10^{-17}$ . If the dissolution of  $\text{Ag}^+$  and  $\text{I}^-$  ions are unequal then at equilibrium the sol will contain charged silver iodide particles. Thus, in the presence of excess iodide ions the particles will be negatively charged and with excess silver ions positively charged particles will be obtained.

### 2.2.4 Isomorphous substitution

The replacement of one atom by another of similar size in a crystal lattice is termed isomorphous substitution. Clays are naturally occurring colloidal particles of size less than  $2 \mu\text{m}$ . The building blocks used to construct the wide variety of clays found in nature are silica and alumina sheets. The silica sheets consist of linked silicon–oxygen tetrahedra with four oxygens surrounding each central-silicon. The alumina sheets are formed from octahedra, each central-aluminium being surrounded by six oxygens or hydroxyls. In simple clays like kaolinite, a sheet of alumina octahedra shares apical oxygens with a sheet of silica tetrahedra. The tetrahedra–octahedra bilayers are then bound together by a combination of van der Waals forces and hydrogen bonding. Clays become charged as a result of amorphous substitution. The substitution of  $\text{Si}^{4+}$  by  $\text{Al}^{3+}$  in the tetrahedral layers or of  $\text{Al}^{3+}$  by  $\text{Mg}^{2+}$ ,  $\text{Zn}^{2+}$  or  $\text{Fe}^{2+}$  in the octahedral layers leads to a net negative charge.

### 2.2.5 Potential determining ions

When discussing the influence of ions on colloid stability it is important to differentiate between inert ions and potential-determining ions. *Potential-determining ions are species which by virtue of their electron distribution between the solid and liquid phases determine the difference in Galvani potential between these phases* (IUPAC). The *Galvani potential*, sometimes termed the *inner potential*, of a phase is the work done in taking a test charge from infinity to a point inside the phase. In the case of silver iodide sols, described above, it is apparent that the surface charge density, and hence the Galvani potential, changes on addition of silver or iodide ions. Hence, for this system  $\text{Ag}^+$  and  $\text{I}^-$  ions are *potential-determining ions*. It is of note that with respect to metal oxides protons are potential-determining ions and thus a change in pH will result in a change in the surface charge. Inert ions do not change the charge density at the surface of particles but may influence the interfacial potential difference by virtue of their local distribution.

## 2.3 The Distribution of Inert Ions at a Charged Interface

If a positively charged surface is placed in an electrolyte containing inert ions then simple electrostatics indicates that cations will be repelled from and anions attracted to the interface. Electroneutrality will be attained when the electrolyte layer near the interface possesses a

net negative charge of equal magnitude to the surface charge of the solid material. This charged atmosphere of electrolyte, whose form will be dependent on both concentration and speciation, will influence the potential drop across the solid/liquid interface.

### 2.3.1 The mercury/electrolyte interface

The relationship between surface charge, interfacial potential difference and electrolyte composition can be experimentally determined for metal/electrolyte interfaces. Hence, though in the majority of colloidal systems the dispersed material is insulating, when developing models of ion distribution at charged interfaces it is instructive to focus on metal/electrolyte interfaces. Historically, the dropping mercury electrode in aqueous electrolyte has been considered as a model metal/electrolyte system, the interface between the two liquids being atomically smooth and the continual renewal of surface minimising contamination. Further, and perhaps most importantly, the mercury/aqueous electrolyte interface is ideally polarisable, i.e. charge carriers are not transferred across the interface, over a wide range of potential differences. The experimental variable for a mercury electrode immersed in electrolyte is the applied potential  $E$  and the measured quantities are surface tension  $\gamma$  and differential capacitance  $C_d$ . The relationships between these experimental parameters and interfacial potential difference and surface charge are detailed below.

The two-electrode assembly that is typically employed for studying the relationship between surface charge and interfacial potential is shown in Figure 2.2, where  $\phi_i$  is the Galvani potential of phase  $i$ . The experimentally applied potential,  $E(\phi_{\text{Cu}_1} - \phi_{\text{Cu}_2})$ , is partitioned across all the interfaces of the system,

$$E = (\phi_{\text{Cu}_1} - \phi_{\text{Hg}}) + (\phi_{\text{Hg}} - \phi_s) + (\phi_s - \phi_{s,\text{ref}}) + (\phi_{s,\text{ref}} - \phi_{\text{ref}}) + (\phi_{\text{ref}} - \phi_{\text{Cu}_2}) \quad (2.1)$$

Careful experimental design ensures that as the applied potential is altered the potential drop across the reference electrode remains constant, and changes to both liquid junction potentials and metal junction potentials are negligible. Hence, for the simple experimental apparatus displayed in Figure 2.2 a change to the applied potential of magnitude  $\Delta E$  results in an equivalent change in the Galvani potential difference across the mercury/electrolyte interface, i.e.

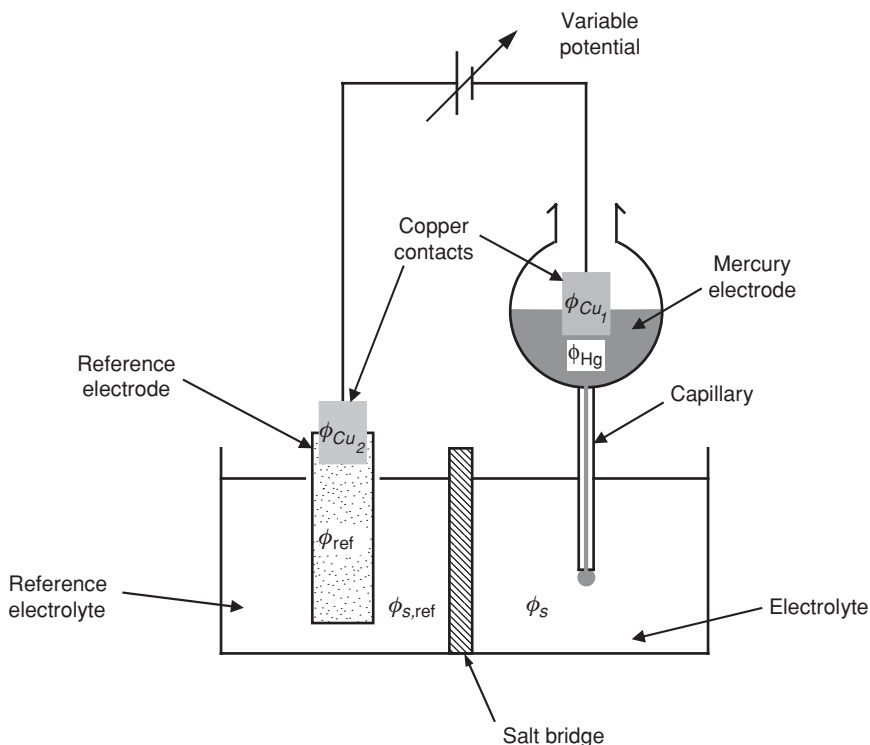
$$\Delta E = \Delta(\phi_{\text{Hg}} - \phi_s) = \Delta\phi_{m/s} \quad (2.2)$$

The Galvani potential difference between the metal and the bulk solution  $\phi_{m/s}$  is the work done in moving unit charge from the solution to a point just inside the metal. When discussing the distribution of ions in the vicinity of the interface it is the Volta potential difference between the electrode surface and bulk electrolyte,  $\psi_{m/s}$ , that is of importance.  $\psi_{m/s}$  is the work done in moving unit charge from the bulk solution to a point just outside the metal. The Volta and Galvani potential differences are related by the expression

$$\phi_{m/s} = \psi_{m/s} + \chi \quad (2.3)$$

where  $\chi$  is the *surface potential* that arises from the work done in transferring the test charge across the surface dipole layer. If it is assumed that  $\chi$  is independent of the applied potential  $E$  then

$$\Delta E = \Delta\phi_{m/s} = \Delta\psi_{m/s} \quad (2.4)$$



**Figure 2.2** The dropping mercury electrode.

A more detailed discussion of the relationship between Volta and Galvani potentials can be found in [2].

The surface charge at a particular applied potential can be obtained by measuring either the surface tension of the mercury drop or the integrated capacitance. For the dropping mercury electrode system shown in Figure 2.2 the surface tension can be calculated from the time required for a drop of maximum size to form,  $t_{\max}$ . Immediately prior to the instant at which the mercury drop detaches from the capillary the forces of gravity and surface tension are equal, i.e.

$$mg t_{\max} = 2\pi r_c \gamma \quad (2.5)$$

where  $m$  is the mass flow rate of mercury,  $r_c$  is the radius of the capillary and  $\gamma$  the mercury surface tension.  $\gamma$  is related to the surface charge density of the metal,  $\sigma_m$ , by the Lippman equation

$$\sigma_m = -\frac{d\gamma}{dE} \quad (2.6)$$

which is derived by equating the work done in charging the electrode,  $\sigma dE$ , to the change in surface free energy,  $-d\gamma$ . When determining the surface charge by impedance methods a hanging mercury drop electrode may be employed. For an ideally polarisable system, in which the solution resistance is negligible, the AC charging current  $i_{CC}$  when a low



amplitude ( $<10$  mV) sinusoidal modulating potential is applied,  $E_{AC} \sin(\omega t)$ , is given by the expression

$$i_{CC} = \omega C_d E_{AC} \cos(\omega t) \quad (2.7)$$

where  $C_d$  is the differential capacitance of the system,

$$C_d = \frac{d\sigma}{dE} \quad (2.8)$$

at the applied DC potential. To determine the surface charge at a potential  $E$  it is necessary to integrate the differential capacitance between the point of zero charge (pzc) and  $E$ . Combining Equations (2.6) and (2.8) gives the following relationships between surface tension, surface charge and differential capacitance:

$$-\frac{d^2\gamma}{dE^2} = \frac{d\sigma}{dE} = C_d \quad (2.9)$$

For a charged surface in a uniform medium simple electrostatics indicates that  $\psi_{m/s}$  is directly proportional to the charge on the metal,  $\sigma_m$ . From Equation (2.4)  $d\psi = dE$ , hence for a mercury electrode in a medium of uniform dielectric a plot of  $\gamma$  against  $E$  is a parabola whilst that of  $C_d$  against  $E$  is a horizontal line, as shown by the dashed lines in Figure 2.3.

For a mercury drop electrode immersed in a  $1.0 \text{ mol dm}^{-3}$  sodium chloride solution the experimentally determined surface charge as a function of potential is shown in Figure 2.3. It is clear that the electrolyte does not act as a uniform medium. This is due to the fact that the mobile anions and cations are non-uniformly distributed in the vicinity of the interface. Figure 2.4 shows electrocapillary and capacitance data for a range of electrolytes. The key points to note from the experimental data are:

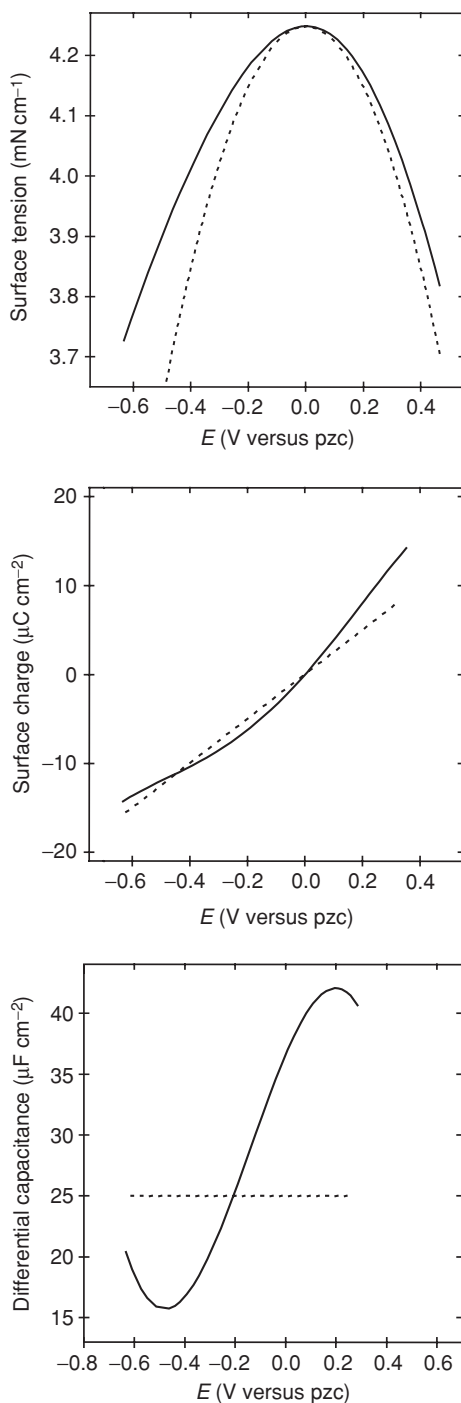
- (a) at low electrolyte concentration the capacitance displays a minimum at the pzc
- (b) at high electrolyte concentrations or high potential  $C_d$  is almost independent of potential
- (c) positive of the pzc the surface tension at a particular potential varies with ion type.

A model of the charged surface/electrolyte interface must account for the above observations. In arriving at a model that satisfies these criteria the chronological development of double layer theory is followed.

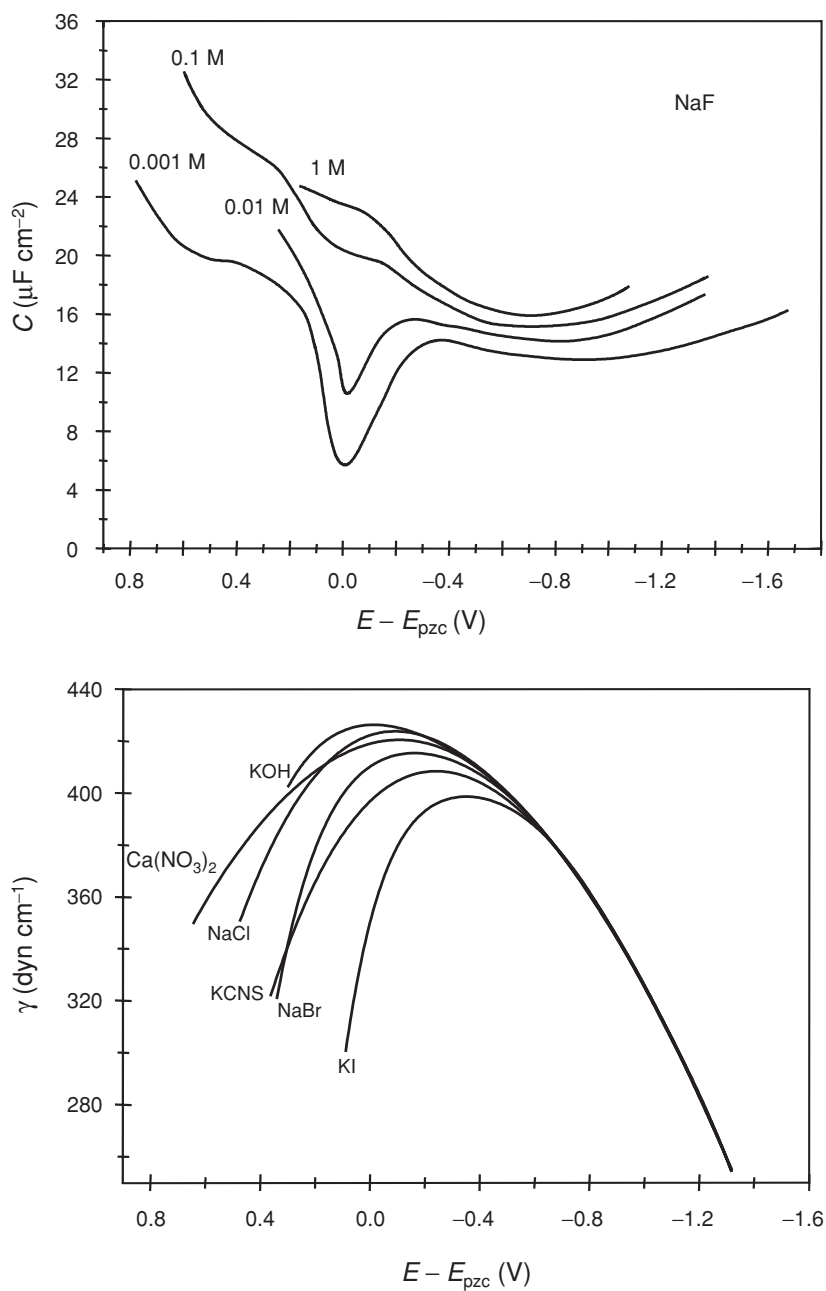
### 2.3.2 The Helmholtz model

A simple model of a charged metal interface immersed in electrolyte was postulated by Helmholtz. He suggested that the surface charge is counterbalanced by a layer of oppositely charged ions whose centre lies a distance  $d$  from the interface, where  $d$  is the radius of the hydrated ions. A schematic of the Helmholtz model, in which the potential is dropped across a single layer of ions, is shown in Figure 2.5. The two layers of opposing charge may be viewed as two plates of a capacitor of capacitance  $C_H$ ,

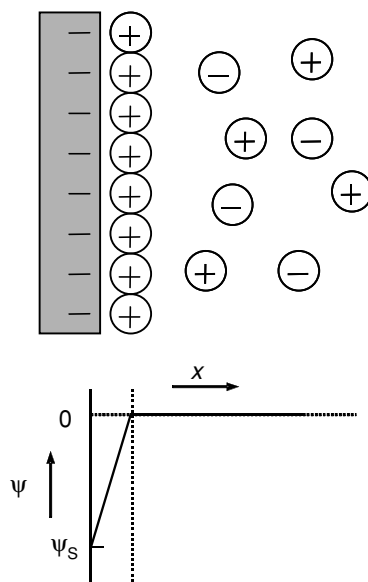
$$C_H = \frac{\epsilon\epsilon_0}{d} \quad (2.10)$$



**Figure 2.3** The (a) surface tension, (b) surface charge and (c) differential capacitance versus  $E$  plots for (solid line) a mercury drop electrode immersed in a  $1.0 \text{ mol dm}^{-3} \text{ NaCl}$  solution and (dashed line) a mercury drop in a uniform medium. (The experimental data are reproduced with permission from [3]).



**Figure 2.4** Plots of (a) differential capacitance and (b) surface tension versus applied potential for various electrolytes (reproduced with permission from [3]).

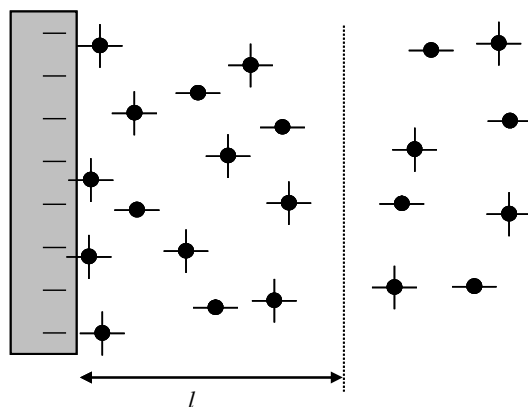


**Figure 2.5** Schematic of the electrolyte near the interface in the Helmholtz model.

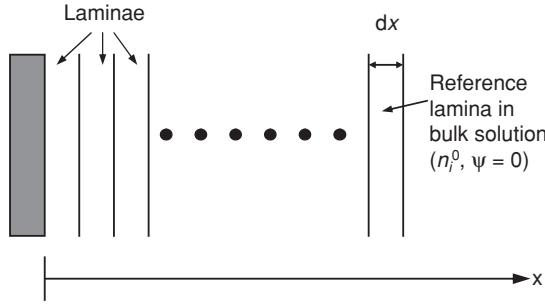
where  $\varepsilon$  is the relative permittivity of the solvent and  $\varepsilon_0$  the permittivity of free space. As  $C_H$  is independent of potential the model fails to predict the experimentally observed minimum in the differential capacitance versus applied potential curve that is observed at low electrolyte concentrations.

### 2.3.3 Gouy–Chapman theory

Gouy and Chapman, working independently, developed a model in which the surface charge of the solid is balanced by a diffuse layer of ions that contains an excess of ions of opposite charge to the surface, as shown in Figure 2.6. In, what is now termed, Gouy–Chapman



**Figure 2.6** Schematic of the electrolyte near the interface in the Gouy–Chapman model.



**Figure 2.7** Schematic of the laminae near the interface in the Gouy–Chapman model.

theory the ions are treated as point charges. Parallel to the electrode surface, where the Volta potential is  $\psi_s$ , the electrolyte is divided into laminae each of thickness  $dx$ , as displayed in Figure 2.7. At large  $x$ , i.e. in the bulk of the solution, the Volta potential is zero and the concentration of the  $i$ th species, of charge  $z_i e$ , is  $c_i^b$ . In a lamina distance  $x$  from the interface the Volta potential is  $\psi(x)$  and the concentration is  $c_i^x$ . The work done in moving an ion from the bulk to a lamina distance  $x$  from the interface is  $z_i e \psi(x)$ . Hence the Boltzmann distribution indicates that

$$c_i^x = c_i^b \exp\left(-\frac{z_i e \psi(x)}{kT}\right) \quad (2.11)$$

The Poisson equation relates the local electric charge density  $\rho_e(x)$ , which may be expressed as  $\sum_i z_i e c_i^x$ , to the local Volta potential gradient:

$$\rho_e(x) = -\frac{\varepsilon \varepsilon_0}{N_A} \left( \frac{d^2 \psi}{dx^2} \right) \quad (2.12)$$

Combining Equations (2.11) and (2.12) yields the Poisson–Boltzmann equation

$$\sum_i z_i e c_i^b \exp\left(-\frac{z_i e \psi}{kT}\right) = -\frac{\varepsilon \varepsilon_0}{N_A} \left( \frac{d^2 \psi}{dx^2} \right) \quad (2.13)$$

If it is assumed that the Volta potential gradient is zero in bulk solution, then for a symmetric electrolyte ( $z_{\text{anions}} = z_{\text{cations}} = z$ ) of bulk concentration  $c^b$  the solution to the Poisson–Boltzmann equation is

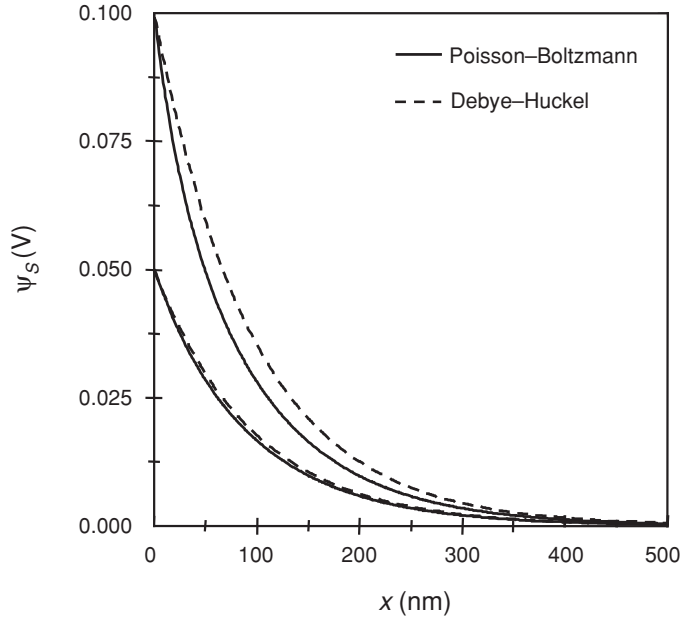
$$\psi(x) = \frac{2kT}{ze} \ln \left( \frac{1 + Z \exp(-\kappa x)}{1 - Z \exp(-\kappa x)} \right) \quad (2.14)$$

where

$$Z = \frac{\exp\left(\frac{ze\psi_s}{2kT}\right) - 1}{\exp\left(\frac{ze\psi_s}{2kT}\right) + 1}$$

and

$$\kappa = \left( \frac{2z^2 e^2 N_A c^b}{\varepsilon \varepsilon_0 kT} \right)^{\frac{1}{2}}$$



**Figure 2.8** Comparison of the potential decay curves for the exact (solid lines) and approximate (dashed lines) solutions to the Poisson-Boltzmann equation.

The charge of the diffuse layer balances the charge of the metal surface, hence

$$\sigma_M = -\sigma_S = -\int_0^\infty \rho_e dx = \frac{2c^b ze N_A}{\kappa} \sinh\left(\frac{ze\psi_S}{2kT}\right) \quad (2.15)$$

and

$$C_d = \frac{d\sigma_M}{d\psi_S} = \frac{2c^b z^2 e^2 N_A}{\kappa kT} \cosh\left(\frac{ze\psi_S}{2kT}\right) \quad (2.16)$$

In the limit  $ze\psi_S \ll kT$ , the Debye-Hückel approximation  $Z \rightarrow \frac{ze\psi_S}{4kT}$  and the Volta potential at a distance  $x$  from the surface tends to  $\psi_S \exp(-\kappa x)$ . The full analytical solution, Equation (2.14), and the result of employing the Debye-Hückel approximation are compared in Figure 2.8.

It is clear that for the surface potentials typically encountered in colloid science the Debye-Hückel approximation gives a realistic potential decay profile. The Debye-Hückel approximation predicts that the surface potential decays exponentially

$$\psi(x) = \psi_S \exp(-\kappa x) \quad (2.17)$$

$\kappa$  has a dimension of  $(\text{length})^{-1}$  and  $1/\kappa$ , the distance at which the Volta potential has fallen to 37% of its magnitude at the surface, is referred to as the diffuse layer thickness or Debye length. For a  $z:z$  aqueous electrolyte at 25°C

$$\kappa = (3.29 \times 10^9) z(c^b)^{\frac{1}{2}} \quad (2.18)$$

**Table 2.1** The dependence of the Debye length on electrolyte concentration

$C^b$ (mol dm <sup>-3</sup> )	Debye length for $z^+ : z^-$ electrolyte (nm)			
	1:1	1:2/2:1	2:2	1:3/3:1
$10^{-1}$	1	0.6	0.5	0.4
$10^{-2}$	3	1.8	1.5	1.2
$10^{-3}$	10	5.6	4.8	3.9
$10^{-4}$	30	18	15	12
$10^{-5}$	100			

In Table 2.1 the Debye lengths at 25°C for aqueous electrolytes of differing stoichiometry are provided. It is apparent that the diffuse layer thickness may vary from a fraction of a nanometres to 100 nm. Expansion of Equations (2.15) and (2.16) in the Debye–Huckel regime yields

$$\sigma_M = \frac{c^b z^2 e^2 N_A}{\kappa k T} \psi_s \quad (2.19)$$

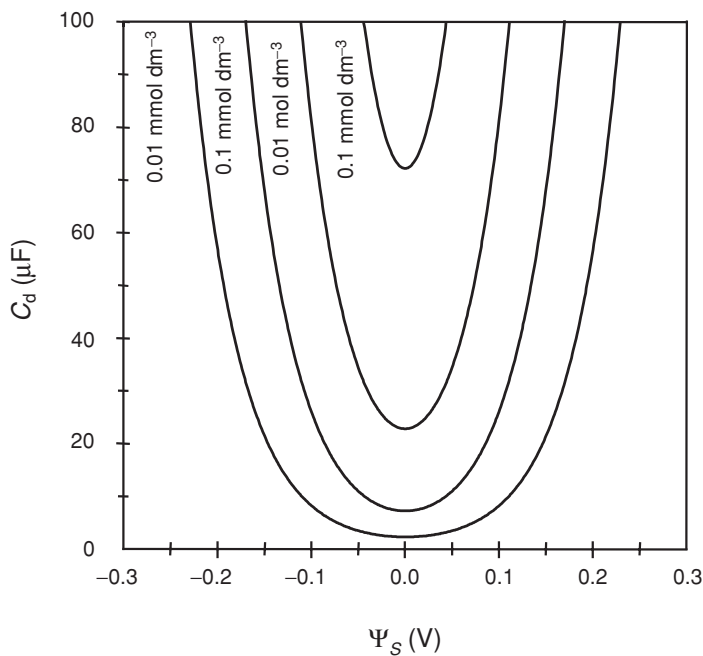
and

$$C_d = \frac{2c^b z^2 e^2 N_A}{\kappa k T} \left[ 1 + \frac{1}{2} \left( \frac{ze\psi_s}{2kT} \right)^2 \right] \quad (2.20)$$

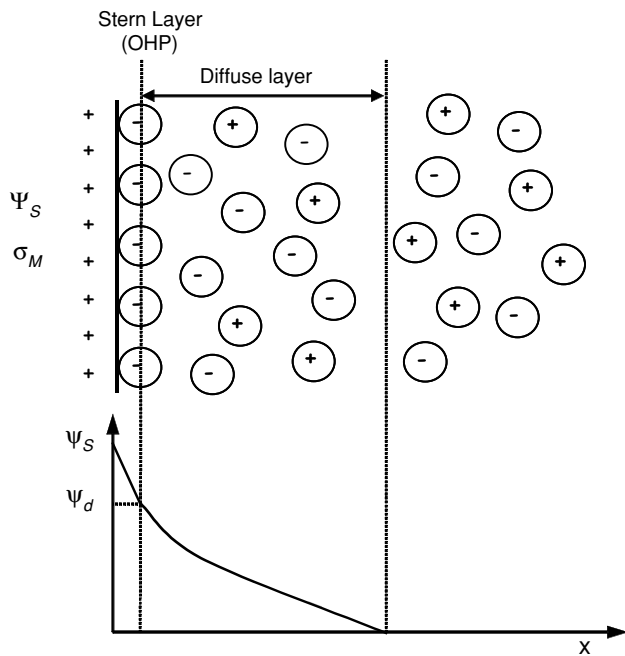
i.e. near the point of zero charge, plots of surface charge and differential capacitance versus applied potential are linear and parabolic respectively. The capacitance versus surface potential predicted by the Gouy–Chapman theory is displayed in Figure 2.9. The plots show a minimum in capacitance at the point of zero charge, and the depth of this minimum varies with the electrolyte concentration. Thus, near the pzc the Gouy–Chapman model correctly predicts the  $C_d$  versus  $E$  behaviour observed in experiments. However, the model fails to account for the concentration independent  $C_d$  observed at high and low potentials. Fatally the differential capacitance goes to infinity at high Volta potential, this breakdown of the model stems from the fact that the point charges may attain unrealistically high concentrations and approach closer to the interface than an ionic radius.

### 2.3.4 The Stern modification

By combining the Helmholtz and Gouy–Chapman models and eliminating the assumption that ions may be treated as point charges Stern developed further the model of the electrode/electrolyte interface. In the Stern theory the counter charge exists in two distinct regions, as shown in Figure 2.10: an inner layer of ions ‘stuck’ to the electrode surface and a diffuse layer of ions. The inner layer lies between the electrode surface and the *outer Helmholtz plane* (OHP), the plane that defines the closest approach of the centre of a hydrated ion to the electrode. The potential at the OHP is  $\psi_d$  and the potential drop across



**Figure 2.9** Differential capacitance versus potential plots for the Gouy–Chapman theory at differing electrolyte concentrations.



**Figure 2.10** Schematic of the electrolyte near the interface in the Stern model.



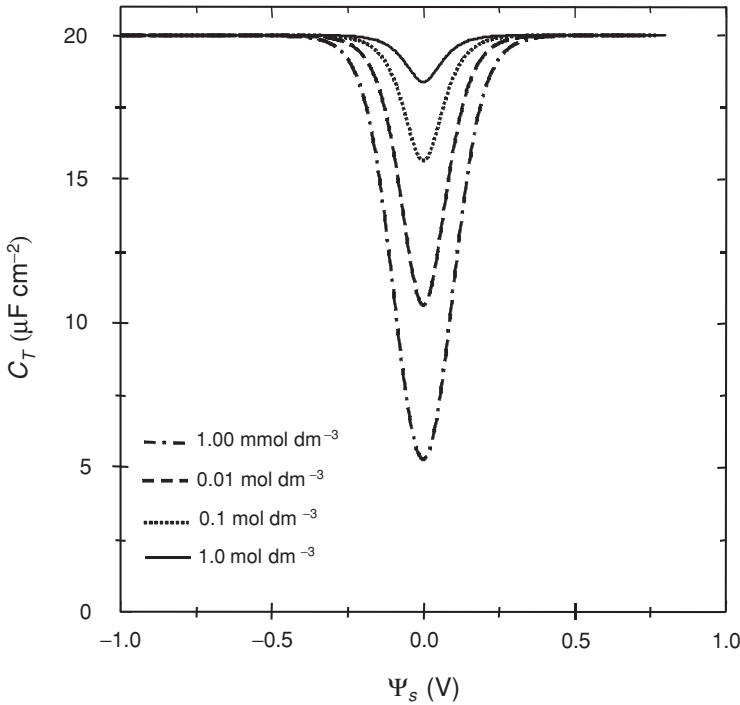
the inner layer ( $\psi_S - \psi_d$ ) is assumed to be linear. Hence, the inner layer may be treated as a parallel plate capacitor of magnitude  $C_i$ :

$$C_i = \frac{\sigma_i}{\psi_S - \psi_d} \quad (2.21)$$

where  $\sigma_i$  is the charge ‘stuck’ at the OHP. As in the Helmholtz model, discussed above, the capacitance is independent of applied potential. The capacitance of the outer diffuse region, where the potential  $\psi_d$  is dropped, is given by the expression

$$C_d = \frac{2c^b z^2 e^2}{\kappa N_A k T} \cosh\left(\frac{ze\psi_d}{2kT}\right) \quad (2.22)$$

The Stern theory considers the interface to consist of two regions of potential drop in series, hence the total capacitance,  $C_T = C_i C_d / C_i + C_d$ , will be dominated by the smallest component. Whilst  $C_i$  is constant,  $C_d$  varies with potential and electrolyte concentration.  $C_T$  as a function of  $\psi_S$  is plotted in Figure 2.11. The Stern model reproduces the essential elements on the experimental capacitance versus potential data discussed above, i.e. a parabolic response near the pzc and a constant response at high and low potentials. To generate all the features of the experimental plots displayed in Figure 2.4, one requires detailed consideration of how the solvent dipoles re-orient as the potential is changed, see for example [2] and [4].



**Figure 2.11** The total capacitance  $C_T$ , as predicted using the Stern model, as a function of  $\psi_S$ .

### 2.3.5 Specific adsorption

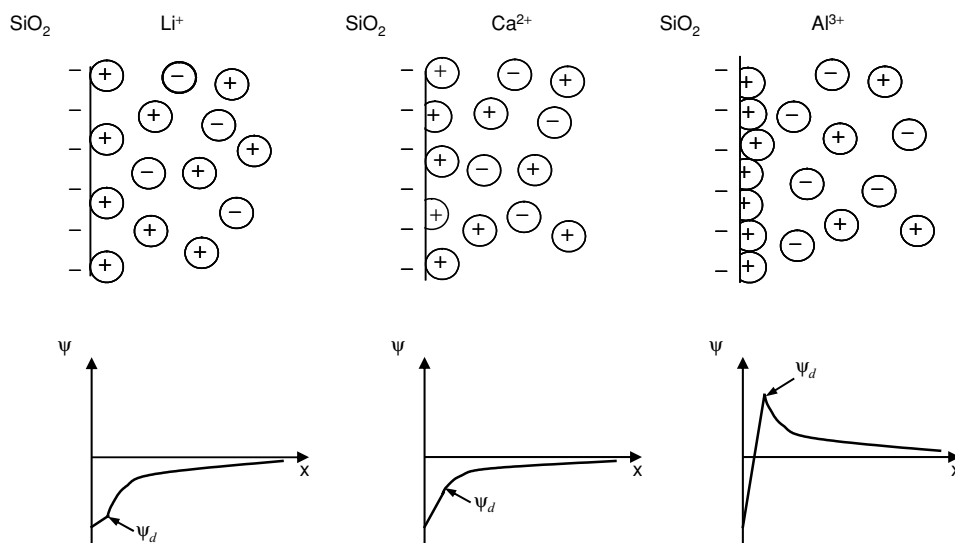
In developing the Stern model the ions that were 'stuck' to the interface were placed at the OHP. The OHP was described as being the plane of closest approach of hydrated ions. However, some ions, as a result of favourable interactions with the surface, may lose all or some of their water of hydration and hence move closer to the interface than the OHP. This process is described as specific adsorption. The average distance of closest approach of the specifically adsorbed ions is termed the *inner Helmholtz plane* (IHP).

On a mercury electrode sodium and fluoride ions are not specifically adsorbed. Hence, sodium fluoride is the electrolyte of choice in experiments to verify models of the electrode–electrolyte interface. On the mercury electrode other anions are specifically adsorbed. The fact that the experimental electrocapillary response at potentials positive of the pzc is electrolyte specific, Figure 2.4, reflects the fact that the anions of the different electrolytes are specifically adsorbed to differing extents.

Schematics of cation adsorption at a silica interface are shown in Figure 2.12. It is of note that as silica adsorption of  $\text{Al}^{3+}$  ions is so favourable that overcompensation of the surface charge occurs and even though the silica is negatively charged, it is necessary for the diffuse layer to contain an excess of negative ions.

### 2.3.6 Inter-particle forces

In obtaining a model of the ion distribution at the interface between a charged phase and electrolyte, a Hg electrode has been considered. We now turn our attention back to charged colloidal particles. In a charged colloid there will be a non-uniform distribution of counter-ions; with specifically adsorbed ions, ions 'stuck' at the OHP and a diffuse layer in which



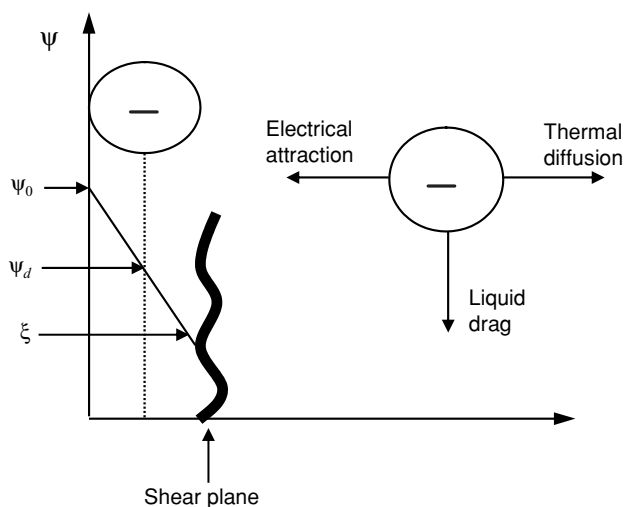
**Figure 2.12** Schematics of the  $\text{SiO}_2$  electrolyte interface in the presence of specifically adsorbing ions.

counter-ions are in excess. The fact that the ionic atmosphere counterbalances the colloid charge begs the question, 'how does the charging of the particles infer stability on a colloidal suspension?' This issue will be discussed in detail in Chapter 3. Here we note that as the particles approach each other, repulsive interactions occur as their diffuse layers begin to overlap. The distribution of ions in the diffuse layer depends on the potential at the OHP,  $\psi_d$ , and the bulk concentration of electrolyte  $c^b$ . This suggests that it is these two parameters that must be measured and controlled in order to formulate stable suspensions. The remainder of this chapter is concerned with methods of approximating  $\psi_d$ .

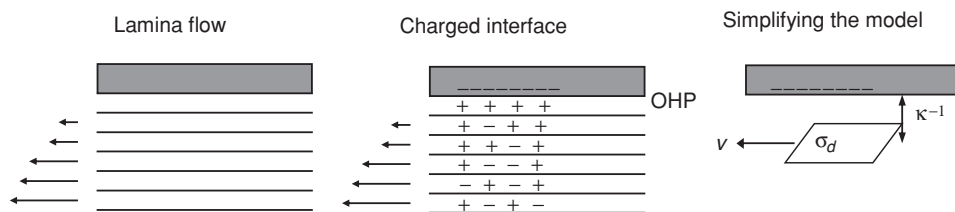
## 2.4 Electrokinetic Properties

So far in describing ion distribution at charged interfaces it has been assumed that both bulk phases are static. As a result of the surface charge, interesting effects occur when one or both of the bulk phases are set in motion. The study of these phenomena is termed *electrokinetics*. In an electrokinetic experiment the solid phase is static and the electrolyte flows, the electrolyte is static and the solid moves or both phases are in motion. There are two experimental strategies in electrokinetic experiments, either the motion is controlled and an electrical property is measured, e.g. *streaming current* and *streaming potential* measurement, or the electrical field is controlled and motion monitored, e.g. *electro-osmotic*, *electrophoretic* and *electrosonic* experiments.

When discussing the relative motion of the two phases in electrokinetics it is useful to define a *shear plane*, the effective location of the solid/liquid interface. Above, the ions whose centre lies on the OHP were described as being 'stuck' to the interface. Under flow conditions these ions remain in contact with the charged phase and the distance of the shear plane from the phase boundary is a hydrated ion diameter, as shown in Figure 2.13. The *zeta potential*  $\zeta$  is the potential at the shear plane and may be experimentally determined



**Figure 2.13** An illustration of the shear plane.



**Figure 2.14** Charge transport under lamina flow conditions.

using electrokinetic methods. It is often assumed, even though the shear plane is at least a hydrated ion diameter from the interface and the outer Helmholtz plane is a hydrated ion radius, that  $\zeta$  is equivalent to  $\psi_d$ .

### 2.4.1 Electrolyte flow

The Reynolds number,  $R_e$ , for the flow of electrolyte, of density  $\rho$  and viscosity  $\eta$ , through a capillary of radius  $a$  is given by the expression

$$R_e = \frac{2\rho v_m a}{\eta} \quad (2.23)$$

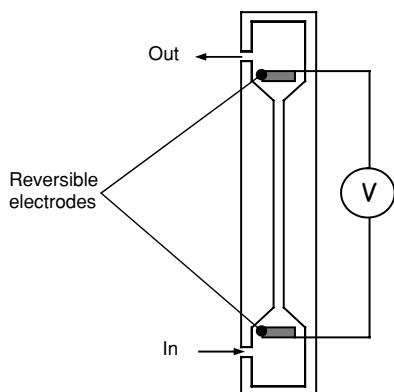
where  $v_m$  is the mean flow rate. If  $R_e$  is less than *ca.* 1000 lamina flow will occur. A parabolic flow profile is established with the laminae in contact with the walls having zero velocity and that at the centre the maximum velocity, as shown in Figure 2.14. If the capillary walls are charged then whilst the ions at the OHP are stationary those in the diffuse layer will flow. As the diffuse layer possesses a net charge this flow leads to a *streaming current*. As the current is proportional to the charge of the diffuse layer, which in turn depends on  $\zeta$ , a study of the streaming current as a function of flow rate allows  $\zeta$  to be determined. To obtain expressions relating streaming current to flow rate the system may be reduced to a single lamina of charge  $\sigma_d$  flowing at a distance  $1/\kappa$  from the interface. Although it is possible to measure streaming currents it is more common to monitor the *streaming potential*. The streaming potential is related to the streaming current by Ohm's law.

### 2.4.2 Streaming potential measurements

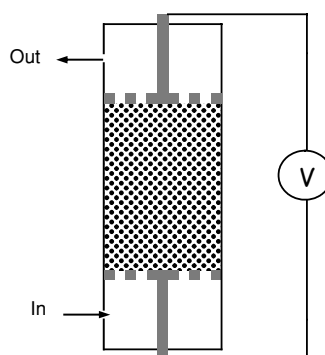
To measure a streaming potential the electrolyte is made to flow past the surface of interest, as shown in Figure 2.15. The streaming potential is measured using polarisable electrodes set perpendicular to the direction of flow. For materials of high dimension a capillary is formed and the solution pumped through the cell. For materials of low dimension, e.g. powders and fibres, a cell consisting of porous electrodes may be employed. In such a cell the streaming potential  $E_s$  is

$$E_s = \frac{\varepsilon \varepsilon_0 \Delta p}{c^b \eta \Lambda} \zeta \quad (2.24)$$

Capillary walls:



Powders or fibres:

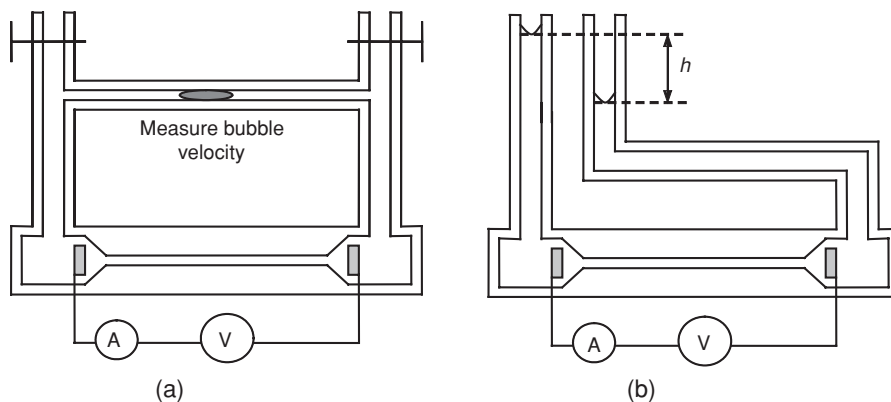


**Figure 2.15** Experimental apparatus for determining streaming potentials.

where  $\Delta p$ , the pressure across the cell, and  $\eta$ , the electrolyte viscosity, describe the hydrodynamics and  $\Lambda$ , the molar conductivity of the solution, stems from application of Ohm's law.

### 2.4.3 Electro-osmosis

If a solution flowing through a capillary leads to a streaming current and a streaming potential then applying a potential  $E$  across a capillary should lead to a flow of solution. The phenomenon of solution flow when a potential is applied is termed *electro-osmosis* and is now often employed to pump solutions in microfluidic systems [5]. Two experiments that employ electro-osmosis to determine the  $\zeta$ -potential of the surface of a capillary are shown in Figure 2.16. In dynamic electro-osmosis measurements, Figure 2.16a, a potential



**Figure 2.16** Experimental apparatus for determining the electro-osmotic pressure.

is applied and the resultant solution flow monitored by measuring the velocity of the bubble. The  $\zeta$ -potential is calculated using the expression

$$\zeta = v_b \frac{a_b^2 i \eta \Lambda}{a_c^2 \varepsilon \varepsilon_0} \quad (2.25)$$

where  $v_b$  is the bubble velocity,  $a_b$  is the radius of the capillary containing the bubble,  $i$  is the current and  $a_c$  is the radius of the capillary across which the potential is applied. In equilibrium electro-osmosis measurements, Figure 2.16b, the pressure required to stop the potential induced flow is measured. The electro-osmotic pressure is related to the  $\zeta$ -potential by the expression

$$\zeta = h \rho g \frac{i}{E^2} \frac{1}{8\pi \varepsilon \varepsilon_0 \Lambda} \quad (2.26)$$

where  $\rho$  is the electrolyte density and  $g$  the acceleration due to gravity.

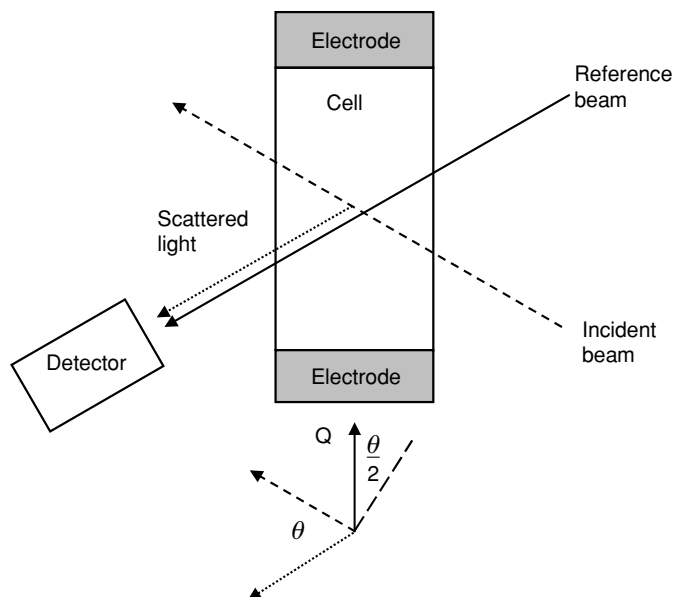
#### 2.4.4 Electrophoresis

In electrophoresis the colloidal particles are the mobile phase and the electrolyte is stationary. Hence, electrophoretic methods are commonly employed when information on the charge of colloidal particles is required. Conceptually it is a very simple technique; optical methods are employed to monitor the velocity of the colloid particles  $v_p$  in an electric field. However, the fact that if the walls of a cell are charged fluid will flow when an electric field is applied means that maintaining a stationary electrolyte layer is experimentally challenging. This problem may be overcome by treating the walls of the electrophoresis cell to prevent charging. Alternatively a closed cell is employed and the optical field is focused on particles in the stationary lamina, which is where the potential induced flow along the cell walls is cancelled by the reverse flow of solution through the centre of the cell. The depth of the stationary solvent lamina is dependent on cell geometry; for a cylindrical cell it is  $0.146d$  and for a rectangular cell  $0.2d$ , where  $d$  is the cell diameter or depth. The velocity of some colloid particles may be measured directly under a microscope. However, techniques in which particle velocity is calculated by determining the change in light scattering when the particles move in an electric field are more commonly employed. Light scattering is discussed in detail in Chapter 12. Here it is considered only in relation to laser Doppler electrophoresis (LDE) and phase analysis light scattering (PALS), techniques that may be employed to determine the velocity of particles in electric fields.

LDE [6] is based on the fact that when laser radiation of frequency  $\nu_0$  is scattered by a particle that is moving in an electric field, the frequency of the scattered light  $\nu_s$  is Doppler shifted. The magnitude of the Doppler shift  $\delta\nu$  is directly proportional to the particle velocity in the direction of the scattering vector. Hence, for quasi-elastic scattering

$$\delta\nu = \frac{2n\nu_p}{\lambda_0} \sin\left(\frac{\theta}{2}\right) \cos\phi \quad (2.27)$$

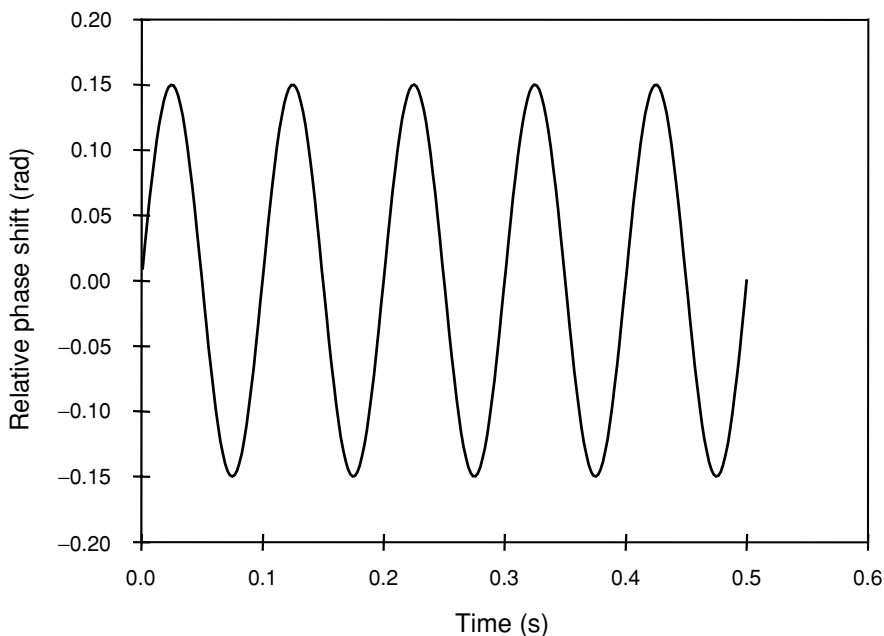
where  $n$  is the refractive index of the suspension medium,  $\lambda_0$  is the wavelength of the laser,  $\theta$  is the scattering angle and  $\phi$  is the angle between the direction of the electric field and the scattering vector. To determine the magnitude of the Doppler shift, which even for a



**Figure 2.17** A schematic of the principal features of the apparatus for laser Doppler electrophoresis.

highly charged mobile particle moving in a high electric field is only a few tens of hertz, the scattered beam is added to a reference beam. A schematic of the apparatus employed in LDE is displayed in Figure 2.17; note that the scattering vector  $Q$  lies along the direction of the electric field, i.e. the angle  $\phi$  is zero. The photocurrent measured at the detector has an AC component at the beat frequency  $|\nu_0 - \nu_s|$ , thus the Doppler shift may be determined and the particle velocity calculated.

The maximum fields employed in LDE experiments are typically of the order of  $10 \text{ V cm}^{-1}$  and are square wave modulated at a frequency of 1 Hz. The form of the field is limited by electrode polarisation and Joule heating of the suspensions. If the charge on colloidal particles is low then their velocity in the maximum applied electric field will be small and, given that the electric field can only be applied for short periods, the Doppler shift in frequency may be immeasurable. Hence, when determining the velocity of particles that exhibit low  $\zeta$ -potential, e.g. in solutions of high ionic strength, suspensions in low polarity solvents or measurements near the isoelectric point, LDE may not have the sensitivity required. Similarly the application of LDE to studies of highly viscous media is limited. To measure the velocity of slow moving particles, PALS [7] was developed. PALS is closely related to LDE but uses a more sophisticated method of comparing the reference and scattered beams and as a result can routinely measure mobilities 1000 times lower than the limit of the LDE technique. A PALS experiment involves phase modulation of the reference beam and modulation of the electric field. In a simplified PALS experiment the reference beam is phase modulated at the frequency of the scattered beam in zero field. The phase shift between the scattered and reference beams is compared. When no field is applied the phase shift will be constant. When a field is applied the phase shift will change with time. Experimentally the relative phase shift, the difference between the phase shift at time  $t$  and the constant phase



**Figure 2.18** The raw data obtained in a PALS experiment when the applied electric field is modulated at a frequency of 10 Hz.

shift in the absence of an electric field, is recorded as a function of time. The relative phase shift can be determined with high accuracy as it may be measured over a high number of cycles. Typically in a PALS experiment a sinusoidal field of frequency in the range 5–60 Hz is applied to the sol and the relative phase shift monitored, see Figure 2.18. The amplitude of the sinusoidally modulated relative phase shift allows the velocity of the particles to be determined. Corrections for the drift velocity of the particles in zero field may be made. Using the PALS technique particle velocities equivalent to Doppler shifts as low as 0.001 Hz may be determined. Hence PALS is the method of choice to determine the velocity of particles in systems where the  $\zeta$ -potential is low or the viscosity is high.

To relate the velocity of the colloid particles to the electric field it is necessary to consider all the forces acting. The motion due to the electric field will be opposed by the viscous drag on the particle, the viscous drag on the ionic atmosphere and the electrostatic force that results from distortion of the diffuse layer. The relative importance of these forces depends on the dimensionless quantity  $\kappa a$  which is the ratio of the radius of curvature of the particle to the double layer thickness. When  $\kappa a$  is small ( $\ll 1$ ) the charged particle may be treated as a point charge and the Huckel equation

$$\zeta = \frac{v_p}{E} \frac{3\eta}{2\epsilon\epsilon_0} \quad (2.28)$$

relates the particle velocity to the  $\zeta$ -potential: generally the Huckel equation is only valid for studies of particles suspended in non-aqueous media of low conductance. When  $\kappa a$  is



large the particle–electrolyte interface may be treated as a flat sheet and the Smoluchowski equation applies:

$$\zeta = \frac{v_p}{E} \frac{\eta}{\varepsilon \varepsilon_0} \quad (2.29)$$

### 2.4.5 Electroacoustic technique

A final method of determining the  $\zeta$ -potential is to use electroacoustics [8]. When a high frequency, approximately 1 MHz, AC field is applied to a colloidal sol, both the particles and the diffuse layer are set in motion. As the inertia of particles differs to that of the diffuse layer the velocity/field transfer function differs for the two components. Pressure waves result and from measurements of the sound the  $\zeta$ -potential may be determined. A notable advantage of the electroacoustic method is that it may be employed with optically dense systems where light scattering techniques are not applicable.

## References

1. Attard, P. (2001) Recent advances in the electric double layer in colloid science. *Curr. Opin. Colloid Interface Sci.*, **6**, 366–371.
2. Bockris, J.O'M. and Reddy, A.K.N. (2000) *Modern Electrochemistry*, 2nd edition. Plenum, New York.
3. Grahame, D.C. (1947) The electrical double layer and the theory of electrocapillarity. *Chem. Rev.*, **41**, 441–501.
4. Delahay, P. (1965) *Double Layer and Electrode Kinetics*. Interscience, New York.
5. Schasfoort, R.B.M., Schlautmann, S., Hendrikse, J. and van der Berg, A. (1999) Field-effect fluid control for microfabricated fluidic networks. *Science*, **286**, 942–945.
6. Uzgiris, E.E. (1981) Laser Doppler methods in electrophoresis. *Prog. Surf. Sci.*, **10**, 53–164.
7. McNeil-Watson, F., Tscharnuter, W. and Miller, J. (1998) A new instrument for the measurement of very small electrophoretic mobilities using phase analysis light scattering (PALS). *Colloids Surf. A*, **140**, 53–57.
8. Hunter, R.J. (1998) Recent developments in the electroacoustic characterisation of colloidal suspensions and emulsions. *Colloids Surf. A*, **141**, 37–65.

# Chapter 3

## Colloid Stability

*John Eastman*

### 3.1 Introduction

One of the important aspects of the study of colloidal dispersions is understanding their stability so that we can manipulate the state of dispersions for particular applications.

Charge stabilisation is one means by which this can be achieved and we can manipulate the stability through changes in the chemical environment such as salt concentration, ion type and pH.

We must understand how this works both in quiescent systems and when external fields, such as gravitational or shear fields, are present.

So what do we mean by stability? This very much depends on the circumstances which are being considered. We can define stability in colloidal systems either in terms of their tendency to aggregate or in terms of their tendency to sediment under the action of gravity.

In this chapter we will focus on the stability to aggregation and look at the factors which control this stability. We will study this by considering the interaction between two representative particles in the system. By considering what happens when two particles come together (during a Brownian collision) we can predict the stability of the whole system by looking at the form of the colloidal pair potential.

### 3.2 The Colloidal Pair Potential

The pair potential is the total potential energy of interaction between two colloidal particles as the separation or distance between them is varied. Formally it is a free energy, and we calculate it by simply summing the various components that we can identify.

The calculation is normally done for two particles in isolation, i.e. at infinite dilution. In a concentrated system (a condensed phase), multi-body interactions should be accounted for, and then we would refer to the potential of mean force. However we get an adequate estimation of the total potential in a concentrated system simply by the summation of the interaction from the nearest neighbours.

It is this interaction energy that governs the stability of colloidal dispersions and which we effectively manipulate whenever we make a change to a formulation.

The components of the pair potential that we are most interested in for charged colloids are those due to the attractive van der Waals forces and the repulsive force between similarly charged particles.

### 3.2.1 Attractive forces

Molecules with a permanent dipole will attract similar molecules as the dipoles align. They will also induce a dipole in an adjacent neutral atom or molecule and cause an attraction. This is relatively easy to understand; however, the motion of the electrons in *any* atom cause rapidly fluctuating dipoles. This leads to the London dispersion interaction as the oscillating dipoles become coupled. Even neutral atoms have a fluctuating dipole due to the motion of the electrons around the nucleus. It is energetically more favourable for adjacent atoms to be oscillating in unison. This is the interaction which we recognise from the non-ideal behaviour of the inert gases.

This interaction is non-directional, so that when large assemblies of atoms are considered, different dipolar orientations do not cancel each other. Colloidal particles are of course large assemblies of atoms and hence the van der Waals forces from the London dispersion interaction act between particles to cause attraction (see Figure 3.1).

The early calculations were due to Hamaker and de Boer [1]. The route is to sum the interaction of one atom in a particle with each atom in the adjacent particle. That interaction is then summed over all the atoms in the first particle.

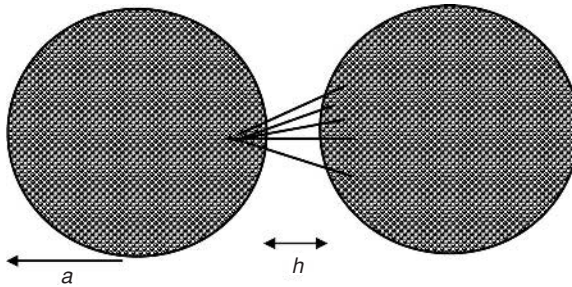
The result is a long range interaction – much longer range than the interaction between two isolated atoms. The range of the interaction is comparable with the radii of colloidal particles.

The attractive potential energy is directly proportional to a particle radius ( $a$ ), a material constant – the Hamaker [1] constant ( $A$ ) – and is inversely proportional to distance of separation ( $h$ ):

$$V_A = -\frac{A}{12} \left[ \frac{1}{x(x+2)} + \frac{1}{(x+1)^2} + 2 \ln \frac{x(x+2)}{(x+1)^2} \right], \quad \text{where } x = \frac{h}{2a} \quad (3.1)$$

When the particle separation is small ( $h \ll 2a$ ) this reduces to a simple form of

$$V_A = -\frac{Aa}{12h} \quad (3.2)$$



**Figure 3.1** London forces between atoms in two adjacent colloidal particles.

**Table 3.1** Hamaker constants for various materials

Particles	Hamaker constant ( $\text{J}/10^{-20}$ )	Media	Hamaker constant ( $\text{J}/10^{-20}$ )
Poly(tetrafluoroethylene)	3.8	Water	3.7
Poly(methyl methacrylate)	7.1	Pentane	3.8
Poly(styrene)	7.8	Ethanol	4.2
Silica (fused)	6.5	Decane	4.8
Titanium dioxide	19.5	Hexadecane	5.1
Metals (Au, Ag, Pt, etc.)	$\sim 40$	Cyclohexane	5.2

The Hamaker constant is a function of both the electronic polarisability and the density of the material. When particles are immersed in a medium the attraction between particles is weakened as there is an attraction with the medium also. The combined or composite Hamaker constant ( $A$ ) can be estimated as the geometric mean of that of the particle ( $A_{\text{particle}}$ ) and that of the medium ( $A_{\text{medium}}$ ) with respect to their values in vacuum, and it is this that should be used in the calculation of the attractive potential:

$$A = (\sqrt{A_{\text{particle}}} - \sqrt{A_{\text{medium}}})^2 \quad (3.3)$$

Hamaker constants have values in the range of  $10^{-20}$  J, and a selection of values are given in Table 3.1.

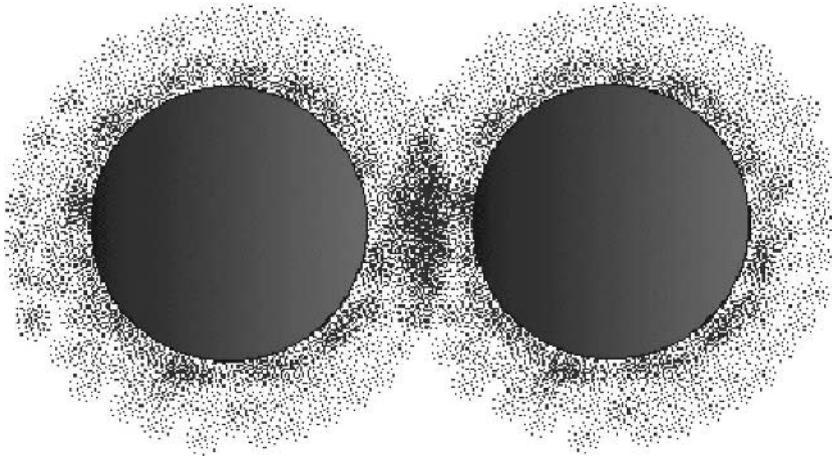
### 3.2.2 Electrostatic repulsion

Electrical repulsion is an important stabilising mechanism for particles dispersed in aqueous solutions or moderate polarity liquids like ethylene glycol.

The diffuse part of the electrical double layer extends in solution over distances characterised by the Debye length ( $1/\kappa$ ). In practice we use the experimentally accessible  $\zeta$ -potential (see Section 3.3.3) as a measure of the electrical potential at the Stern layer. The rate of decay of this potential is governed by the reciprocal of the Debye length and is commonly referred to as the double layer thickness (see Table 3.2). This defines the extent to which the ionic atmosphere, which is different from the bulk ionic medium, extends from the particle surface.

**Table 3.2** The extent of the double layer thickness as a function of electrolyte concentration

NaCl Concentration (mM)	Double layer thickness (nm)
30	2
10	3
1	10
0.1	30



**Figure 3.2** The overlap of electrical double layers in adjacent particles.

When two particles approach each other the ionic atmospheres overlap (see Figure 3.2) and the local ion concentration midway between the particles can be estimated by summing the contributions from each particle. The difference in this local mid-point ion concentration and that in the bulk results in an osmotic pressure acting to force the particles apart. Integration of this force with respect to distance gives us the energy.

When two charged particles come together there are two extreme cases which we can envisage. If the ion adsorption equilibrium is maintained then either the surface charge remains constant and the surface potential compensates (constant charge) or the surface potential remains constant and the surface charge density changes to compensate (constant potential). Hogg *et al.* [2] derived expressions which enable us to calculate the interaction between non-identical spheres under both constant charge and constant potential conditions.

$$V_R^\psi = \frac{\varepsilon a_1 a_2 (\psi_{0_1}^2 + \psi_{0_2}^2)}{4(a_1 + a_2)} \left[ \frac{2\psi_{0_1} \psi_{0_2}}{\psi_{0_1}^2 + \psi_{0_2}^2} \ln \left( \frac{1 + \exp(-\kappa h)}{1 - \exp(-\kappa h)} \right) + \ln(1 - \exp(-2\kappa h)) \right] \quad (3.4)$$

$$V_R^\sigma = \frac{\varepsilon a_1 a_2 (\psi_{0_1}^2 + \psi_{0_2}^2)}{4(a_1 + a_2)} \left[ \frac{2\psi_{0_1} \psi_{0_2}}{\psi_{0_1}^2 + \psi_{0_2}^2} \ln \left( \frac{1 + \exp(-\kappa h)}{1 - \exp(-\kappa h)} \right) - \ln(1 - \exp(-2\kappa h)) \right] \quad (3.5)$$

These reduce to the basic expressions

$$V_R^\psi = \frac{\varepsilon a \psi_0^2}{2} \ln(1 + \exp(-\kappa h)) \quad (3.6)$$

$$V_R^\sigma = -\frac{\varepsilon a \psi_0^2}{2} \ln(1 - \exp(-\kappa h)) \quad (3.7)$$

for the interaction between identical particles with a radius  $a$ . The expressions are valid in the regime where  $\kappa a$ , the product of the Debye constant and the particle radius, is greater than 10.

For conditions where  $\kappa a$  is less than 3 the general expression is

$$V_R = 2\pi\epsilon a\psi_\delta^2 \exp(-\kappa h) \quad (3.8)$$

### 3.2.3 Effect of particle concentration

Whenever we add charged colloidal particles to a liquid we do two things:

- add counter-ions with each particle
- reduce the solution volume available to the ions.

Both of these factors become important as the particle concentration increases and when the background electrolyte concentration is low. We can expand the expression for  $\kappa$  from Equation (2.18), Chapter 2, to take account of these extra effects. The expanded expression is

$$\kappa^2 = \frac{e^2 z^2}{\epsilon k_B T} \frac{2n_0 + \frac{3|\sigma_\delta|\phi}{ae}}{1 - \phi} \quad (3.9)$$

where  $z$  is the counter-ion valency,  $n_0$  the concentration of counter-ions in solution (added electrolyte),  $a$  is the particle radius and  $e$  is the formal charge on an electron.

We recognise the first part of the expression from Equation (2.18) but now we have the expression  $\frac{3|\sigma_\delta|\phi}{ae}$  which takes into account the counter-ions which are carried by the particle through the surface charge density  $\sigma_\delta$ . We can see that this expression becomes important when the particles are small and the volume fraction  $\phi$  and surface charge density are high. The denominator of  $(1 - \phi)$  takes into account the volume taken up by the particles in the dispersion.

This effect is only important when the background electrolyte levels are low and where we have high concentrations of small highly charged particles. We can see the effect in Figure 3.3 which shows the effect for different background electrolytes as a function of particle volume fraction for 85 nm radius particles with a surface charge density of  $0.15 \mu\text{C cm}^{-2}$ .

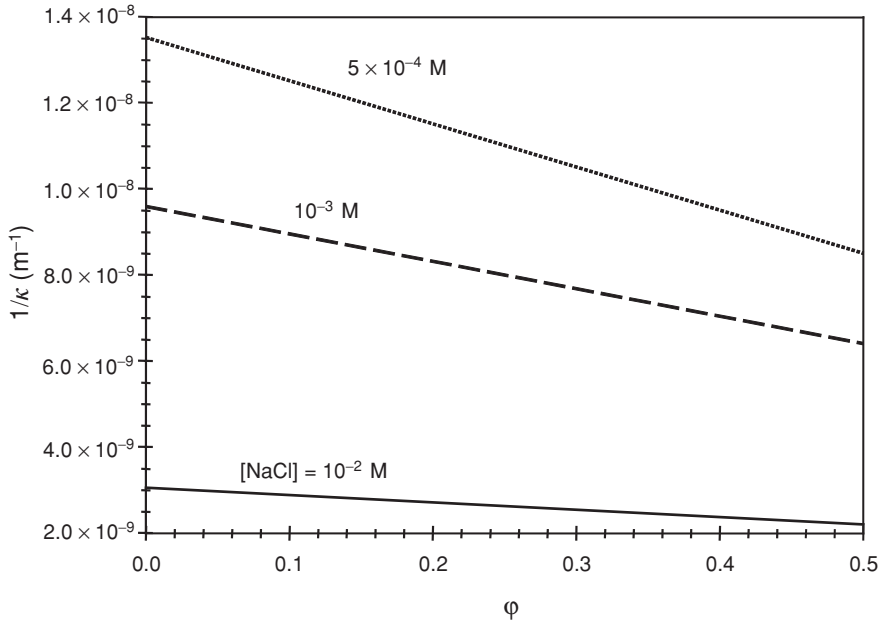
### 3.2.4 Total potential

The linear addition of the electrostatic and dispersion potentials is the basis of the DLVO [3, 4] theory for colloid stability. When we add the attractive potential to the repulsive electrostatic potential we have the typical curve for charge stabilised colloidal particles:

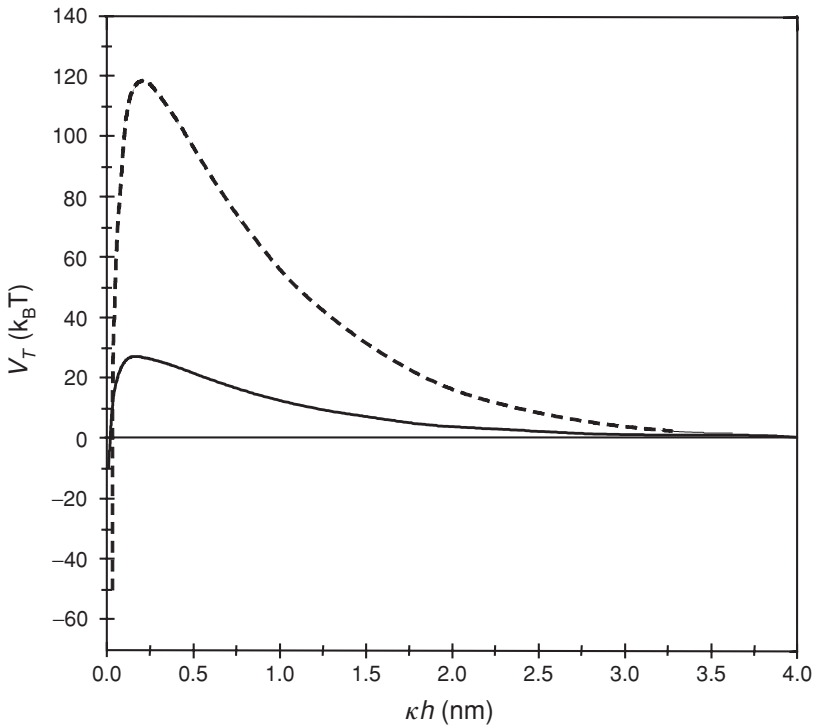
$$V_T = V_A + V_R \quad (3.10)$$

This curve has a number of interesting and important features. The shape of the curve is the consequence of the addition of the exponential decay of the repulsive term and the more steeply decaying one-over-distance relationship of the attractive term.

This linear superposition leads to a maximum in the curve as seen in Figure 3.4 and is known as the primary maximum. It is this maximum in the pair potential which provides



**Figure 3.3** The effect of the counter-ions associated with the particle surface on the Debye length  $1/\kappa$ .



**Figure 3.4** Examples of a total interaction potential curve for two charge stabilised systems.

the mechanism for stability of charged colloidal particles. It creates an effective activation energy for aggregation. As two particles come together they must collide with sufficient energy to overcome the barrier provided by the primary maximum.

It is important to realise that this barrier to aggregation only provides kinetic stability to a dispersion. The thermodynamic drive is towards an aggregated, phase separated state. We can say that the larger the barrier the longer the system will remain stable.

Note that the potential is plotted in units of  $k_B T$ . These thermal energy units help us relate the height of the maximum with the energy of a Brownian collision, which will be of the order of  $1.5k_B T$ .

Therefore, in order to pose a suitable barrier to aggregation this primary maximum must be at least  $1.5k_B T$ . In practice we need to manipulate the system so that the primary maximum is at least  $20k_B T$  in order to achieve a level of stability which can be relied upon over an extended period of time.

### 3.3 Criteria for Stability

We need to define how the various factors that we put into our systems affect the stability so that we can define threshold values.

We need to consider

- the effects of ion type and concentration
- the value of the  $\zeta$ -potential
- the effect of particle size.

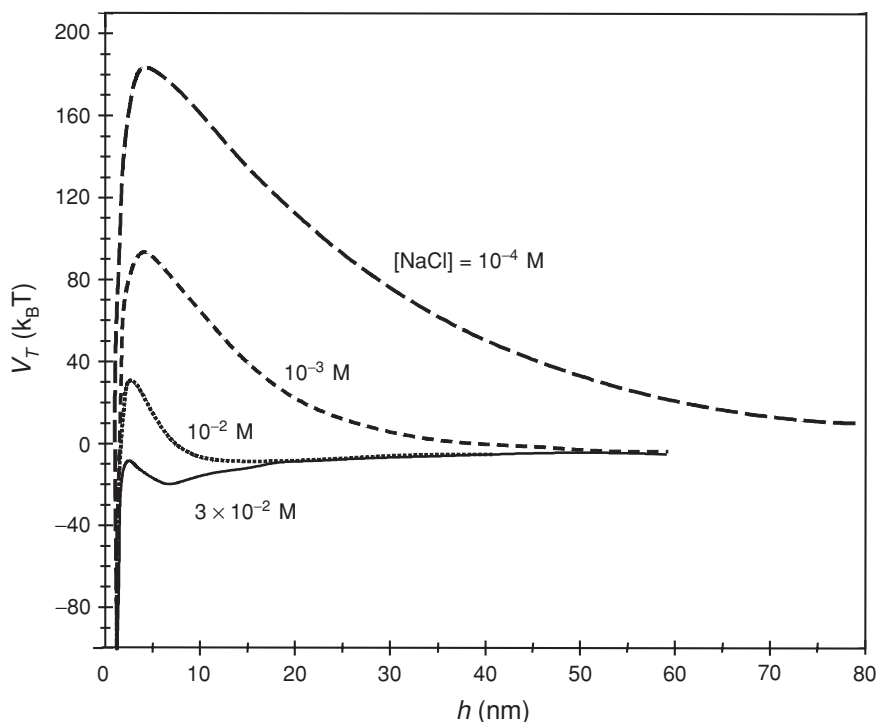
#### 3.3.1 Salt concentration

The example given in Figure 3.5 shows the curve for titanium dioxide particles of 100 nm radius and with a  $\zeta$ -potential of  $-50$  mV at different concentrations of sodium chloride.

The points to note are that:

- In each case there is a steep rise to the primary maximum.
- At larger distances there is a long repulsive tail, most notable at lower electrolyte concentrations.
- The range of the tail reduces as the electrolyte concentration increases (in line with the decrease in the Debye length).
- The height of the maximum decreases with increasing electrolyte concentration.
- At some point (in this case around  $10^{-2}$  M NaCl) a significant energy minimum develops since the van der Waals dispersion term is insensitive to the electrolyte changes. The attractive minimum is known as the secondary minimum.
- As the primary maximum falls to just a few  $k_B T$ , a significant fraction of colliding particles will collide with at least that energy and stick.
- As the primary maximum falls to below zero (above  $3 \times 10^{-2}$  M NaCl in this case), all collisions will lead to aggregation as there is no barrier.





**Figure 3.5** The effect of salt concentration on the shape of the total interaction potential curve.

### 3.3.2 Counter-ion valency

The counter-ions are the dominant ions in the Stern and diffuse layers and hence the stability is more sensitive to the counter-ion type than the co-ion type. The valency of the counter-ions is in fact of major importance in determining the stability of charged colloids.

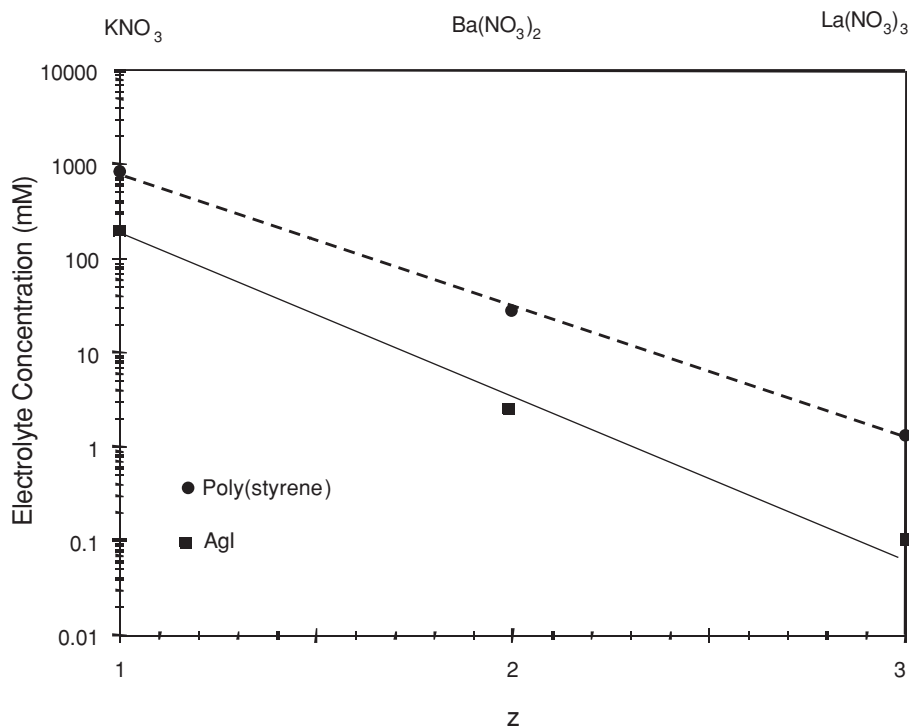
We can estimate a critical coagulation concentration (ccc) from the pair potential exercise by taking the condition that when there is no primary maximum barrier, the inter-particle force is also zero and so the coagulation of the particles will be diffusion controlled (all collisions result in coagulation).

Early observations noted a 6th power dependence of  $z$  on the ccc, and this was formalised in the Shultz–Hardy rule [5–7]:

$$\text{ccc} \propto \frac{1}{z^n} \quad (3.11)$$

- $n = 6$  for high potentials (unusual for coagulation due to ion adsorption)
- $n = 2$  for low potentials (the more common occurrence).

With many systems the adsorption of ions and resultant decrease in the Stern potential and hence  $\zeta$ -potential (see also Section 2.1 of Chapter 2) result in a power less than 6; however, a trivalent counter-ion such as  $\text{Al}^{3+}$  is six times as effective a coagulant as  $\text{Na}^+$ .



**Figure 3.6** Critical coagulation concentrations for two different colloidal systems using three different electrolytes.

Figure 3.6 shows some experimentally determined values of the ccc for two types of dispersions. Both systems show the marked sensitivity to the counter-ion valency predicted by the Shultz–Hardy rule. The more-dense AgI particles have a higher Hamaker constant than the polystyrene and hence are more easily aggregated and have lower values of the ccc at each valency of the counter-ion. The line plotted shows the expected 6th power dependence.

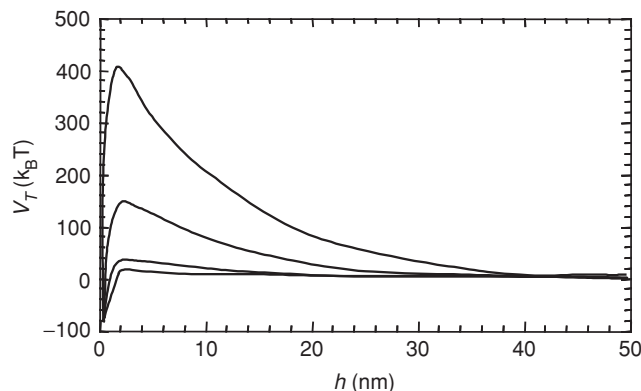
It is important to note that the pH of the systems containing the trivalent ion was held at pH = 4. It is common for trivalent ions such as  $\text{Al}^{3+}$ ,  $\text{Fe}^{3+}$  and  $\text{La}^{3+}$  to form large hydrated complexes with higher valencies than 3, at pH values greater than 4. Hence at pH 7 or 8, aluminium chloride is an even more effective coagulant than at pH 3 or 4.

### 3.3.3 $\zeta$ -potential

The  $\zeta$ -potential is a characteristic that is often experimentally accessible. We can expect it to be close to the value of the Stern potential and hence it is often used in the calculation of the pair potential. It may be possible to determine the surface charge density through a titration procedure and to then calculate the Stern potential, but this is experimentally laborious.

In the expression for the electrostatic repulsion the surface potential appears as the square, and so it is a key parameter in estimating the primary maximum:

$$V_R = 2\pi\epsilon a\psi_\delta^2 \exp(-\kappa h) \quad (3.12)$$



**Figure 3.7** The effect of  $\zeta$ -potential on the shape of the total interaction potential curve for a polystyrene latex. From the bottom to top the  $\zeta$ -potentials are  $-20$  mV,  $-25$  mV,  $-50$  mV and  $-80$  mV.

In the example given in Figure 3.7 for polymer latex particles in 1 mM NaCl, we need a value in excess of 20 mV to produce a stable dispersion.

When the  $\zeta$ -potential is  $-25$  mV the value of  $V_{\max}$  is  $\sim 40k_B T$ . When we double that to  $-50$  mV we can see that  $V_{\max} \sim 160k_B T$ . So as expected, since  $V_R$  is related to the square of the  $\zeta$ -potential, a doubling of the  $\zeta$ -potential leads to a quadrupling of the value of  $V_{\max}$ .

In this example, when the  $\zeta$ -potential reduces to less than  $-20$  mV the value of  $V_{\max}$  drops below  $20k_B T$  and significant aggregation will occur.

### 3.3.4 Particle size

Both the attractive and repulsive contributions are proportional to the particle radius. At small sizes the value of  $V_T$  is directly proportional to the particle size. However, at large sizes the value of  $V_T$  has a more complicated variation.

In all cases a larger particle radius leads to a higher energy barrier; in other words, electrostatic stability increases with increasing particle radius (all other factors remaining constant). For small particle sizes ( $<100$  nm radius) the primary maximum is directly proportional to the radius. However, the relationship breaks down at larger sizes and the height of the primary maximum increases at a lower rate.

The shape of the curve of the attractive interaction is the important point here. Another feature of this is that for particles with large radii, the attraction often dominates again at long range giving rise to a secondary minimum at distances of the order of 5–10 nm. This attraction is manifest as weak but reversible aggregation.

We can make a distinction here between two types of aggregation. Coagulation is the rapid aggregation that happens in the absence of a primary maximum and leads to a strong irreversible aggregated structure. Flocculation is a reversible aggregation that occurs in a secondary minimum as described. Flocculation is reversible on the addition of energy to the system, usually the application of a shear field by shaking, stirring or other mechanical processes.

### 3.4 Kinetics of Coagulation

The rate of coagulation is used either directly or indirectly to determine the ccc. For example if the rate is monitored, we find that it increases as the electrolyte concentration is increased until it reaches a plateau value. If we just add particles to different electrolyte solutions in a series of tubes, then we can check for the onset of aggregation after a fixed time. This may be 5 min, or we may choose a little longer, but it is still the faster rate that we notice.

As we have already established that electrostatically stabilised dispersions are kinetically stable and not thermodynamically stable, the key factor is the kinetics. If the rate is so slow that we do not detect a significant change during our period of use, we would consider that to be adequately stable.

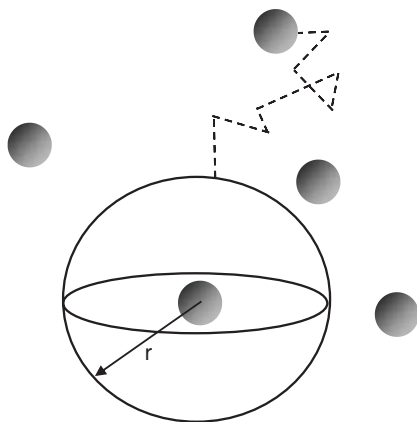
#### 3.4.1 Diffusion limited rapid coagulation

Recall that the diffusion constant is in terms of the flux through a unit area per second ( $J_p$ ):

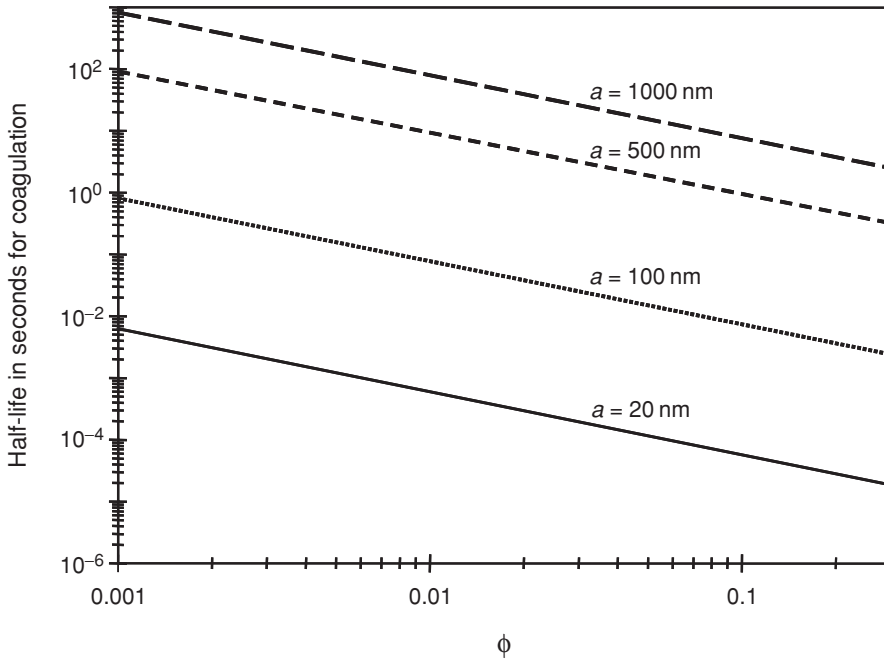
$$J_p = D4\pi r^2 \frac{dN}{dr} \quad (3.13)$$

We can calculate the flow through a spherical surface around a reference particle (see Figure 3.8). This gives us a differential equation that is easily solved to give the number of collisions with that reference particle. Of course each particle is itself such a particle and so the total number of collisions is that for one particle multiplied by the total particle number. We must divide by 2 as particle A colliding with particle B is the same collision as particle B colliding with particle A.

We can allow for the fact that all particles are in motion by using the sum of the diffusion constants of the two colliding particles. As we are assuming them all to be of the same size, we multiply by 2. As each collision results in coagulation the coagulation rate is simply the collision rate.



**Figure 3.8** Theoretical spherical surface of influence surrounding a reference particle.



**Figure 3.9** Half-life for coagulation of a series of dispersions of different particle sizes as a function of volume fraction. From top to bottom, 1000 nm, 500 nm, 100 nm and 20 nm radii.

We may write the rate constant in terms of the diffusion constant and the particle radius. The diffusion constant is

$$k_D = 8\pi Da \quad (3.14)$$

and the half-life is

$$t_{\frac{1}{2}} = \frac{3\eta}{4k_B T N_p} \quad (3.15)$$

where  $\eta$  is the viscosity. We may write the half-life for the second order reaction in terms of the reciprocal of the particle number. Figure 3.9 shows how the half-life decreases with size and concentration.

### 3.4.2 Interaction limited coagulation

When there is an energy barrier to prevent the particles coming together the rate is slowed, as only a fraction of the particles collide with sufficient energy to exceed the height of the barrier and stick. This is known as reaction limited aggregation with a rate constant  $k_R$  in contrast to the diffusion limited rate constant  $k_D$ .

Fuchs [8] defined the stability ratio  $W$  as the ratio of the rate constants so that the higher the stability ratio the slower the rate:

$$W = \frac{k_D}{k_R} \quad (3.16)$$

To a good approximation the reaction limited diffusion rate constant is proportional to the Boltzmann factor which gives the fraction of particles at any instant with energy in excess of the primary maximum. The rate of aggregation drops rapidly with the increase in the energy barrier, so that  $10\text{--}20k_B T$  gives us reasonable kinetic stability.

Overbeek [4] has shown that a reasonable approximation to the stability ratio is obtained using the value of the primary maximum:

$$W = \frac{1}{2\kappa a} \exp\left(\frac{V_{\max}}{k_B T}\right) \quad (3.17)$$

from which we can show that

$$k_R \approx 16\pi\kappa Da^2 \exp\left(\frac{-V_{\max}}{k_B T}\right) \quad (3.18)$$

### 3.4.3 Experimental determination of the ccc

The ccc is the salt concentration at which there is a change from aggregation which is limited by the presence of a primary maximum to aggregation which has no barrier. At this point there is a distinct change in the coagulation rate. Measurement of the coagulation rate can be made in two ways

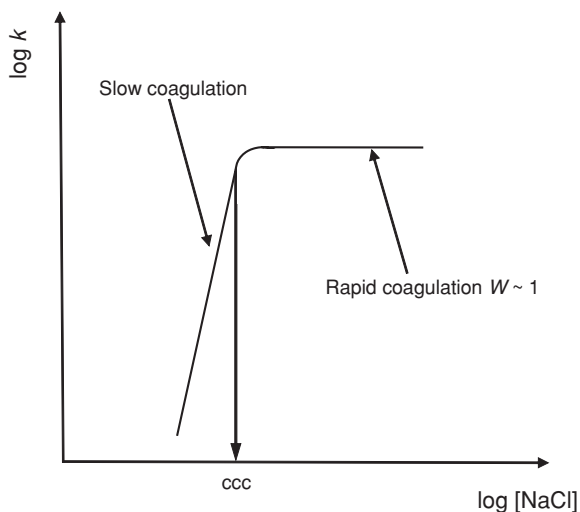
- directly – by measuring the number of particles as a function of time by particle counting (best for large particles)
- indirectly – by light scattering (best for small particles).

In each case we are observing the change in the number of aggregates with time. This tells us about the loss of primary particles from the system.

The simplest experiment is to make up a series of tubes with different electrolyte concentrations and observe at which concentration aggregation becomes apparent after, say, 5 or 10 min. It can be quantified if the tubes are lightly centrifuged and a spectrophotometer is used to measure the per cent transmission of the supernatant as a measure of the number of particles in suspension.

More precise determination can be made if the rate of aggregation itself is measured. At higher electrolyte concentrations the rate increases to the plateau value (as seen in Figure 3.10) representing the fast or diffusion limited rate.

Note that the analysis of the rate constant in terms of the diffusion of single particles is strictly the initial rate. As we progress into the coagulation process the particle number changes and the mechanism changes to one where the large much less mobile aggregates get larger by adding singlets. Accurately describing the rates can then be quite complex.



**Figure 3.10** Rate of coagulation against salt concentration indicating the ccc.

### 3.5 Conclusions

In this chapter we have explored the basic theory surrounding the stability of systems containing charged colloidal particles. The balance between the van der Waals interaction and the repulsion between the electrical double layers surrounding charged particles can be controlled to provide an energetic barrier to the coagulation of particles.

We have seen that charge stabilisation can only be effective where significant surface charge can be achieved and so is normally limited to systems in polar solvents.

The stability achieved is only a kinetic stability to coagulation and in the absence of any other stabilising mechanism these systems will eventually coagulate. However, in the stability ratio we have a method for evaluating the rate of coagulation compared with the theoretical rate when there is no barrier to aggregation.

### References

1. Hamaker, H.C. (1937) *Physica*, **4**, 1058.
2. Hogg, R., Healey, T.W. and Fuerstenau, D.W. (1966) *Trans. Faraday Soc.*, **62**, 1638.
3. Derjaguin, B.V. and Landau, L. (1941) *Acta Physicochim. (URSS)*, **14**, 633.
4. Verwey, E.J. and Overbeek, J.T.G. (1948) *Theory of the Stability of Lyophobic Colloids*. Elsevier, Amsterdam.
5. Schulze, J. (1882) *pr. Chem.*, **25**, 471.
6. Schulze, J. (1883) *pr. Chem.*, **27**, 320.
7. Hardy, W.B. (1900) *Proc. R. Soc.*, **66a**, 110.
8. Fuchs, N. (1934) *Z. Phys.*, **89**, 736.

## Chapter 4

# Surfactant Aggregation and Adsorption at Interfaces

*Julian Eastoe*

A major group of colloidal systems, also classified as lyophilic, is that of the so-called *association colloids* [1]. These are aggregates of *amphiphilic* (both ‘oil and water-loving’) molecules that associate in a dynamic and thermodynamically driven process that may be simultaneously a molecular solution and a true colloidal system. Such molecules are commonly termed ‘surfactants’, a contraction of the term *surface-active agents*. Surfactants are an important and versatile class of chemicals. Due to their dual nature, they are associated with many useful interfacial phenomena, e.g. wetting, and as such are found in many diverse industrial products and processes.

### 4.1 Characteristic Features of Surfactants

Surface-active agents are organic molecules that when dissolved in a solvent at low concentration have the ability to adsorb (or locate) at interfaces, thereby altering significantly the physical properties of those interfaces. The term ‘interface’ is commonly employed here to describe the boundary in liquid/liquid, solid/liquid and gas/liquid systems, although in the latter case the term ‘surface’ can also be used. This adsorption behaviour can be attributed to the solvent nature and to a chemical structure for surfactants that combine both a polar and a non-polar (amphiphilic) group into a single molecule. To accommodate for their dual nature, amphiphiles therefore ‘sit’ at interfaces so that their lyophobic moiety keeps away from strong solvent interactions, while the lyophilic part remains in solution. Since water is the most common solvent, and is the liquid of most academic and industrial interest, amphiphiles will be described with regard to their ‘hydrophilic’ and ‘hydrophobic’ moieties, or ‘head’ and ‘tail’ respectively.

Adsorption is associated with significant energetic changes since the free energy of a surfactant molecule located at the interface is lower than that of a molecule solubilised in either bulk phase. Accumulation of amphiphiles at the interface (liquid/liquid or gas/liquid) is therefore a spontaneous process and results in a decrease of the interfacial (surface) tension. However, such a definition applies to many substances: medium- or long-chain alcohols are surface active (e.g. *n*-hexanol, dodecanol) but these are not considered as surfactants. True surfactants are distinguished by an ability to form oriented monolayers at the interface (here



air/water or oil/water) and, most importantly, self-assembly structures (micelles, vesicles) in bulk phases. They also stand out from the more general class of surface-active agents owing to emulsification, dispersion, wetting, foaming or detergency properties.

Both adsorption and aggregation phenomena result from the hydrophobic effect [2]; i.e. the expulsion of surfactant tails from water. Basically this originates from water–water intermolecular interactions being stronger than those between water–tail. Finally another characteristic of surfactants, when their aqueous concentration exceeds approximately 40%, is an ability to form liquid crystalline phases (or lyotropic mesophases). These systems consist of extended aggregation of surfactant molecules into large organised structures.

Owing to such a versatile phase behaviour and diversity in colloidal structures, surfactants find application in many industrial processes, essentially where high surface areas, modification of the interfacial activity or stability of colloidal systems is required. The variety of surfactants and the synergism offered by mixed-surfactant systems [3] also explains the ever-growing interest in fundamental studies and practical applications. Listing the various physical properties and associated uses of surfactants is beyond the scope of this chapter. However, a few relevant examples are presented in the following section, giving an idea of their widespread industrial use.

## 4.2 Classification and Applications of Surfactants


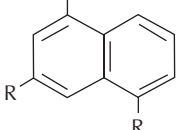
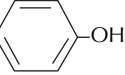
### 4.2.1 Types of surfactants

Numerous variations are possible within the structure of both the head and tail groups of surfactants. The head group can be charged or neutral, small and compact in size or a polymeric chain. The tail group is usually a single or double, straight or branched hydrocarbon chain, but may also be a fluorocarbon or a siloxane, or contain aromatic group(s). Commonly encountered hydrophilic and hydrophobic groups are listed in Tables 4.1 and 4.2 respectively.

**Table 4.1** Common hydrophilic groups found in commercially available surfactants

Class	General structure
Sulfonate	$R-SO_3^- M^+$
Sulfate	$R-OSO_3^- M^+$
Carboxylate	$R-COO^- M^+$
Phosphate	$R-OPO_3^- M^+$
Ammonium	$R_xH_yN^+X^-$ ( $x = 1 - 3$ , $y = 4 - x$ )
Quaternary ammonium	$R_4N^+X^-$
Betaines	$RN^+(CH_3)_2CH_2COO^-$
Sulfobetaines	$RN^+(CH_3)_2CH_2CH_2SO_3^-$
Polyoxyethylene (POE)	$R-OCH_2CH_2(OCH_2CH_2)_nOH$
Polyols	Sucrose, sorbitan, glycerol, ethylene glycol, etc
Polypeptide	$R-NH-CHR-CO-NH-CHR'-CO-\dots-CO_2H$
Polyglycidyl	$R-(OCH_2CH[CH_2OH]CH_2)_n-\dots-OCH_2CH[CH_2OH]CH_2OH$

**Table 4.2** Common hydrophobic groups used in commercially available surfactants

Group	General structure	
Natural fatty acids	$\text{CH}_3(\text{CH}_2)_n$	$n = 12\text{--}18$
Olefins	$\text{CH}_3(\text{CH}_2)_n\text{CH}=\text{CH}_2$	$n = 7\text{--}17$
Alkylbenzenes	$\text{CH}_3(\text{CH}_2)_n\text{CH}_2$ 	$n = 6\text{--}10$ , linear or branched
Alkylaromatics	$\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ 	$n = 1\text{--}2$ for water soluble, $n = 8$ or $9$ for oil soluble surfactants
Alkylphenols	$\text{CH}_3(\text{CH}_2)_n\text{CH}_2$ 	$n = 6\text{--}10$ , linear or branched
Polyoxypropylene	$\text{CH}_3\underset{\text{X}}{\text{CH}}\text{CH}_2\text{O}(\underset{\text{CH}_3}{\text{CH}}\text{CH}_2)_n$	$n$ = degree of oligomerisation, $\text{X}$ = oligomerisation initiator
Fluorocarbons	$\text{CH}_3(\text{CF}_2)_n\text{COOH}$	$n = 4\text{--}8$ , linear or branched, or H-terminated
Silicones	$\text{CH}_3\text{O}(\underset{\text{CH}_3}{\text{Si}}\text{O})_n\text{CH}_3$	

Since the hydrophilic part normally achieves its solubility either by ionic interactions or by hydrogen bonding, the simplest classification is based on surfactant head group type, with further subgroups according to the nature of the lyophobic moiety. Four basic classes therefore emerge as:

- the anionics and cationics, which dissociate in water into two oppositely charged species (the surfactant ion and its counterion)
- the non-ionics, which include a highly polar (non-charged) moiety, such as polyoxyethylene ( $-\text{OCH}_2\text{CH}_2\text{O}-$ ) or polyol groups
- the zwitterionics (or amphoteric), which combine both a positive and a negative group.

With the continuous search for improving surfactant properties, new structures have recently emerged that exhibit interesting synergistic interactions or enhanced surface and aggregation properties. These novel surfactants have attracted much interest, and include the catanionics, bolaforms, gemini (or dimeric) surfactants and polymeric and polymerisable surfactants [4, 5]. Characteristics and typical examples are shown in Table 4.3. Another important driving force for this research is the need for enhanced surfactant biodegradability. In particular for personal care products and household detergents, regulations [6] require high biodegradability and non-toxicity of each component present in the formulation.

A typical example of a double-chain surfactant is sodium bis(2-ethylhexyl)sulfosuccinate, often referred to by its American Cyanamid trade name Aerosol-OT, or AOT. Its chemical

**Table 4.3** Structural features and examples of new surfactant classes

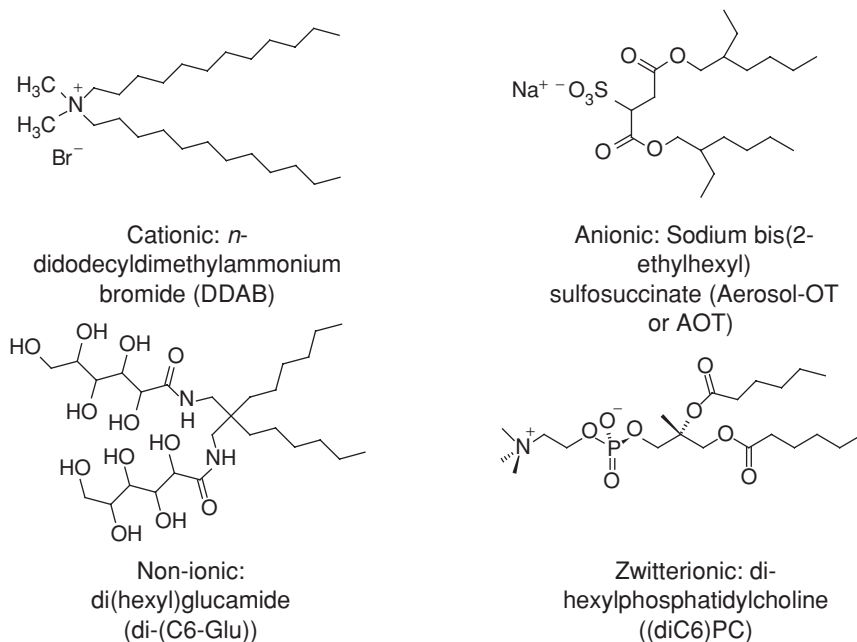
Classes	Structural characteristics	Example
Catanionic	Equimolar mixture of cationic and anionic surfactants (no inorganic counterion)	<i>n</i> -dodecyltrimethylammonium <i>n</i> -dodecylsulfate (DTADS) $C_{12}H_{25}(CH_3)_3N^+ - O_4SC_{12}H_{25}$
Bolaform	Two charged headgroups connected by a long linear polymethylene chain	Hexadecanediy-1,16-bis(trimethylammonium bromide) $Br^-(CH_3)_3N^+-(CH_2)_{16}-N^+(CH_3)_3Br^-$
Gemini (or dimeric)	Two identical surfactants connected by a spacer close to or at the level of the headgroup	Propane-1,13-bis(dodecyltrimethylammonium bromide) $C_3H_6-1,3-bis[(CH_3)_3N^+ C_{12}H_{25}Br^-]$
Polymeric	Polymer with surface active properties	Copolymer of isobutylene and succinic anhydride $H_3C \left( \begin{array}{c} CH_3 \\   \\ -C- \\   \\ CH_3 \end{array} CH_2 \right)_n CH_2 \begin{array}{c} CH \\   \\ CH_2 \\   \\ COOH \end{array} - \overset{O}{\parallel} C - N - CH_2CH_2OH$
Polymerisable	Surfactant that can undergo homopolymerisation or copolymerisation with other components of the system	11-(acryloyloxy) undecyltrimethylammonium bromide $CH_2=CH-C(=O)O-(CH_2)_{10}-N^+(CH_3)_3 Br^-$

structure is illustrated in Figure 4.1, along with other typical double-chain compounds within the four basic surfactant classes.

#### 4.2.2 Surfactant uses and development

Surfactants may be from natural or synthetic sources. The first category includes naturally occurring amphiphiles such as the lipids, which are surfactants based on glycerol and are vital components of the cell membrane. Also in this group are the so-called soaps, the first recognised surfactants [7]. These can be traced back to Egyptian times; by combining animal and vegetable oils with alkaline salts a soap-like material was formed, and this was used for treating skin diseases, as well as for washing. Soaps remained the only source of natural detergents until the early twentieth century, with gradually more varieties becoming available for shaving and shampooing, as well as bathing and laundering. In 1916, in response to a World War I-related shortage of fats for making soap, the first synthetic detergent was developed in Germany. Known today simply as detergents, synthetic detergents are washing and cleaning products obtained from a variety of raw materials.

Nowadays, synthetic surfactants are essential components in many industrial processes and formulations [8–10]. Depending on the precise chemical nature of the product, the



**Figure 4.1** Chemical structure of typical double-chain surfactants.

properties of, for example emulsification, detergency and foaming may be exhibited in varying degree. The number and arrangement of the hydrocarbon groups together with the nature and position of the hydrophilic groups combine to determine the surface-active properties of the molecule. For example C12–C20 is generally regarded as the range covering optimum detergency, whilst wetting and foaming are best achieved with shorter chain lengths. Structure-performance relationships and chemical compatibility are therefore key elements in surfactant-based formulations, so that much research is devoted to this area.

Amongst the different classes of surfactants, anionics are often used in applications, mainly because of the ease and low cost of manufacture. They contain negatively charged head group, e.g. carboxylates ( $-\text{CO}_2^-$ ), used in soaps, sulfate ( $-\text{OSO}_3^-$ ) and sulfonates ( $-\text{SO}_3^-$ ) groups. Their main applications are as detergents, personal care products, emulsifiers and soaps.

Cationics have positively charged head groups – e.g. trimethylammonium ion ( $-\text{N}(\text{CH}_3)_3^+$ ) – and are mainly involved in applications related to their absorption at surfaces. These are generally negatively charged (e.g. metal, plastics, minerals, fibres, hairs and cell membranes) so that they can be modified upon treatment with cationic surfactants. They are therefore used as anticorrosion and antistatic agents, flotation collectors, fabric softeners, hair conditioners and bactericides.

Non-ionics contain groups with a strong affinity for water due to strong dipole–dipole interactions arising from hydrogen bonding, e.g. ethoxylates ( $-(\text{OCH}_2\text{CH}_2)_m\text{OH}$ ). One advantage over ionics is that the length of both the hydrophilic and hydrophobic groups can be varied to obtain maximum efficiency in use. They find applications in low temperature detergents and emulsifiers.

Zwitterionics constitute the smallest surfactant class due to their high cost of manufacture. They are characterised by excellent dermatological properties and skin compatibility. Because of their low eye and skin irritation, common uses are in shampoos and cosmetics.

## 4.3 Adsorption of Surfactants at Interfaces

### 4.3.1 Surface tension and surface activity

Due to the different environment of molecules located at an interface compared to those from either bulk phase, an interface is associated with a surface free energy. At the air–water surface for example, water molecules are subjected to unequal short-range attraction forces and, thus, undergo a net inward pull to the bulk phase. Minimisation of the contact area with the gas phase is therefore a spontaneous process, explaining why drops and bubbles are round. The surface free energy per unit area, defined as *surface tension* ( $\gamma_o$ ), is then the minimum amount of work ( $W_{\min}$ ) required to create new unit area of that interface ( $\Delta A$ ), so  $W_{\min} = \gamma_o \times \Delta A$ . Another, but less intuitive, definition of surface tension is given as the force acting normal to the liquid–gas interface per unit length of the resulting thin film on the surface.

A surface-active agent is therefore a substance that at low concentrations adsorbs, thereby changing the amount of work required to expand that interface. In particular surfactants can significantly reduce interfacial tension due to their dual chemical nature. Considering the air–water boundary, the force driving adsorption is unfavourable hydrophobic interactions within the bulk phase. There, water molecules interact with one another through hydrogen bonding, so the presence of hydrocarbon groups in dissolved amphiphilic molecules causes distortion of this solvent structure apparently increasing the free energy of the system: this is known as the hydrophobic effect [11]. Less work is required to bring a surfactant molecule to the surface than a water molecule, so that migration of the surfactant to the surface is a spontaneous process. At the gas–liquid interface, the result is the creation of new unit area of surface and the formation of an *oriented surfactant monolayer* with the hydrophobic tails pointing out of, and the head group inside, the water phase. The balance against the tendency of the surface to contract under normal surface tension forces causes an increase in the surface (or expanding) pressure  $\pi$ , and therefore a decrease in surface tension  $\gamma$  of the solution. The surface pressure is defined as  $\pi = \gamma_o - \gamma$ , where  $\gamma_o$  is the surface tension of a clean air–water surface.

Depending on the surfactant molecular structure, adsorption takes place over various concentration ranges and rates, but typically, above a well-defined concentration – the critical micelle concentration (CMC) – micellisation or aggregation takes place. At the CMC, the interface is at (near) maximum coverage and to minimise further free energy, molecules begin to aggregate in the bulk phase. Above the CMC, the system then consists of an adsorbed monomolecular layer, free monomers and micellised surfactant in the bulk, with all these three states in equilibrium. The structure and formation of micelles will be briefly described in Section 4.5. Below the CMC, adsorption is a dynamic equilibrium with surfactant molecules perpetually arriving at, and leaving, the surface. Nevertheless, a time-averaged value for the surface concentration can be defined and quantified either directly or indirectly using thermodynamic equations (see Section 4.3.2).

*Dynamic* surface tension – as opposed to the equilibrium quantity – is an important property of surfactant systems as it governs many important industrial and biological applications [12–15]. Examples are printing and coating processes where an equilibrium surface tension is never attained, and a new area of interface is continuously formed. In any surfactant solution, the equilibrium surface tension is not achieved instantaneously and surfactant molecules must first diffuse from the bulk to the surface, then adsorb, whilst also achieving the correct orientation. Therefore, a freshly formed interface of a surfactant solution has a surface tension very close to that of the solvent, and this dynamic surface tension will then decay over a certain period of time to the equilibrium value. This relaxation can range from milliseconds to days depending on the surfactant type and concentration. In order to control this dynamic behaviour, it is necessary to understand the main processes governing transport of surfactant molecules from the bulk to the interface. This area of research therefore attracts much attention and recent developments can be found in [16–18]. However, in the present chapter equilibrium surface tension will always be considered.

### 4.3.2 Surface excess and thermodynamics of adsorption

Following on the formation of an oriented surfactant monolayer, a fundamental associated physical quantity is the *surface excess*. This is defined as the concentration of surfactant molecules in a surface plane, relative to that at a similar plane in the bulk. A common thermodynamic treatment of the variation of surface tension with composition has been derived by Gibbs [19].

An important approximation associated with this Gibbs adsorption equation is the ‘exact’ location of the interface. Consider a surfactant aqueous phase  $\alpha$  in equilibrium with vapour  $\beta$ . The interface is a region of indeterminate thickness  $\tau$  across which the properties of the system vary from values specific to phase  $\alpha$  to those characteristic of  $\beta$ . Since properties within this real interface cannot be well defined, a convenient assumption is to consider a mathematical plane, with zero thickness, so that the properties of  $\alpha$  and  $\beta$  apply right up to that dividing plane positioned at some specific value  $X$ . Figure 4.2 illustrates this idealised system.

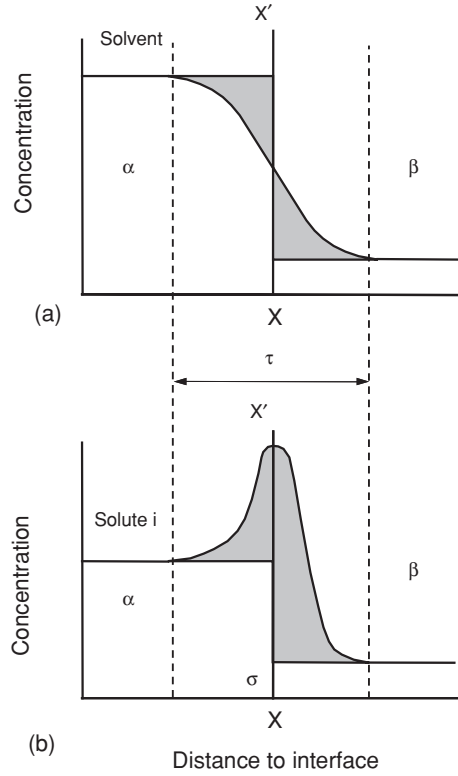
In the definition of the Gibbs dividing surface,  $XX'$  is arbitrarily chosen so that the surface excess adsorption of the solvent is zero. Then the surface excess concentration of component  $i$  is given by

$$\Gamma_i^\sigma = \frac{n_i^\sigma}{A} \quad (4.1)$$

where  $A$  is the interfacial area. The term  $n_i^\sigma$  is the amount of component  $i$  in the surface phase  $\sigma$  over and above that which would have been in the phase  $\sigma$  if the bulk phases  $\alpha$  and  $\beta$  had extended to the surface  $XX'$ , without any change of composition.  $\Gamma_i^\sigma$  may be positive or negative, and its magnitude clearly depends on the location of  $XX'$ .

Now consider the internal energy  $U$  of the total system consisting of the bulk phases  $\alpha$  and  $\beta$

$$\begin{aligned} U &= U^\alpha + U^\beta + U^\sigma \\ U^\alpha &= TS^\alpha - PV^\alpha + \sum_i \mu_i n_i^\alpha \\ U^\beta &= TS^\beta - PV^\beta + \sum_i \mu_i n_i^\beta \end{aligned} \quad (4.2)$$



**Figure 4.2** In the Gibbs approach to defining the surface excess concentration  $\Gamma$ , the Gibbs dividing surface is defined as the plane in which the solvent excess concentration becomes zero (the shaded area is equal on each side of the plane) as in (a). The surface excess of component  $i$  will then be the difference in the concentrations of that component on either side of that plane (the shaded area) (b).

The corresponding expression for the thermodynamic energy of the interfacial region  $\sigma$  is

$$U^\sigma = TS^\sigma + \gamma A + \sum_i \mu_i n_i^\sigma \quad (4.3)$$

For any infinitesimal change in  $T$ ,  $S$ ,  $A$ ,  $\mu$ ,  $n$ , differentiation of Equation (4.3) gives

$$dU^\sigma = TdS^\sigma + S^\sigma dT + \gamma dA + Ad\gamma + \sum_i \mu_i dn_i^\sigma + \sum_i n_i^\sigma d\mu_i \quad (4.4)$$

For a small, isobaric, isothermal, reversible change the differential total internal energy in any bulk phase is

$$dU = TdS - PdV + \sum_i \mu_i dn_i \quad (4.5)$$

Similarly for the differential internal energy in the interfacial region

$$dU^\sigma = TdS^\sigma + \gamma dA + \sum_i \mu_i dn_i^\sigma \quad (4.6)$$

Subtracting Equation (4.6) from (4.4) leads to

$$S^\sigma dT + Ad\gamma + \sum_i n_i^\sigma d\mu_i = 0 \quad (4.7)$$

Then at constant temperature, with the surface excess of component  $i$ ,  $\Gamma_i^\sigma$ , as defined in Equation (4.1), the general form of the Gibbs equation is

$$d\gamma = - \sum_i \Gamma_i^\sigma d\mu_i \quad (4.8)$$

For a simple system consisting of a solvent and a solute, denoted by the subscripts 1 and 2 respectively, then Equation (4.8) reduces to

$$d\gamma = -\Gamma_1^\sigma d\mu_1 - \Gamma_2^\sigma d\mu_2 \quad (4.9)$$

Considering the choice of the Gibbs dividing surface position, i.e. so that  $\Gamma_1^\sigma = 0$ , then Equation (4.9) simplifies to

$$d\gamma = -\Gamma_2^\sigma d\mu_2 \quad (4.10)$$

where  $\Gamma_2^\sigma$  is the solute surface excess concentration.

The chemical potential is given by  $\mu_i = \mu_i^o + RT \ln a_i$  so at constant temperature

$$d\mu_i = \text{const} + RT d \ln a_i \quad (4.11)$$

where  $\mu_i^o$  is the standard chemical potential of component  $i$ .

Therefore applying to Equation (4.10) gives the common form of the Gibbs equation for non-dissociating materials (e.g. non-ionic surfactants)

$$d\gamma = -\Gamma_2^\sigma RT d \ln a_2 \quad (4.12)$$

or

$$\Gamma_2^\sigma = -\frac{1}{RT} \frac{d\gamma}{d \ln a_2} \quad (4.13)$$

For dissociating solutes, such as ionic surfactants of the form  $R^-M^+$  and assuming ideal behaviour below the CMC, Equation (4.12) becomes

$$d\gamma = -\Gamma_R^\sigma d\mu_R - \Gamma_M^\sigma d\mu_M \quad (4.14)$$

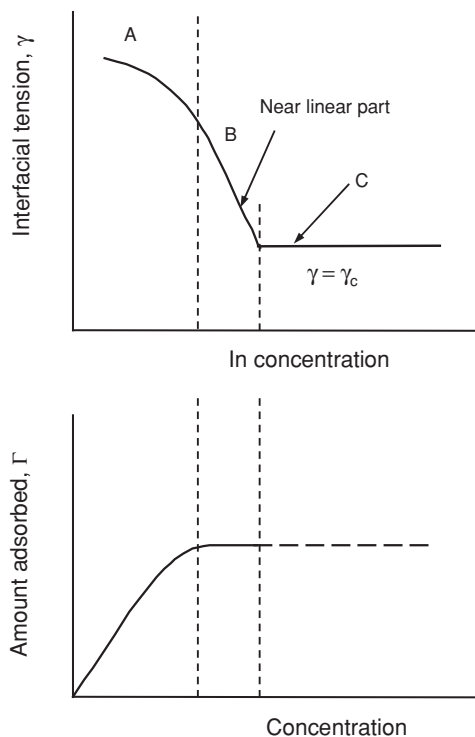
If no electrolyte is added, electroneutrality of the interface requires that  $\Gamma_R^\sigma = \Gamma_M^\sigma$ . Using the mean ionic activities so that  $a_2 = (a_R a_M)^{1/2}$  and substituting in Equation (4.14) gives the Gibbs equation for 1:1 dissociating compounds

$$\Gamma_2^\sigma = -\frac{1}{2RT} \frac{d\gamma}{d \ln a_2} \quad (4.15)$$

If swamping electrolyte is introduced (i.e. sufficient salt to make electrostatic effects unimportant) and the same gegenion  $M^+$  as the surfactant is present, then the activity of  $M^+$  is constant and the pre-factor becomes unity, so that Equation (4.13) is appropriate.

For materials that are strongly adsorbed at an interface such as surfactants, a dramatic reduction in interfacial (surface) tension is observed with small changes in bulk phase concentration. The practical applicability of this relationship is that the relative adsorption





**Figure 4.3** Determination of the interfacial adsorption isotherm from surface tension measurement and the Gibbs adsorption equation.

of a material at an interface, its surface activity, can be determined from the measurement of the interfacial tension as a function of solute concentration. Note that in Equations (4.13) and (4.15), for dilute surfactant systems, the concentration can be substituted for activity without loss of generality.

Figure 4.3 shows a typical decay of surface tension of water on increase in surfactant concentration, and how the Gibbs equation (Equation (4.13) or (4.15)) is used to quantify adsorption at the surface. At low concentrations a gradual decay in surface tension is observed (from the surface tension of pure water i.e.  $72.5 \text{ mN m}^{-1}$  at  $25^\circ\text{C}$ ) corresponding to an increase in the surface excess of component 2 (region A to B). Then at concentrations close to the CMC, the adsorption tends to a limiting value so the surface tension curve may appear to be essentially linear (region B to C). However, in practice, for most surfactants in the pre-CMC region the  $\gamma$ - $\ln c$  is curved so that the local tangent  $-d\gamma/d\ln c$  is proportional to  $\Gamma_2^\sigma$  via Equation (4.13) or (4.15). For single-chain, pure surfactants typical values for  $\Gamma_2^\sigma$  at the CMC are in the range  $2\text{--}4 \times 10^{-6} \text{ mol m}^{-2}$ , with the associated limiting molecular areas being  $0.4\text{--}0.6 \text{ nm}^2$ .

The value for the Gibbs pre-factor in the case of ionic surfactants has been a matter of discussion (e.g. [20–23]). Of particular concern is the question whether, in the case of ionics, complete dissociation occurs giving rise to a pre-factor of 2, or a depletion layer in the sub-surface could be present so that a somewhat lower pre-factor could be expected.

Recent detailed experiments combining tensiometry and neutron reflectivity, which enables direct measurement of the surface excess (as detailed in Chapter 12), have confirmed the use of a pre-factor of 2 in the case of ionics [24].

Although the Gibbs equation is the most commonly used mathematical relation for adsorption at liquid–liquid and liquid–gas interfaces, other adsorption isotherms have been proposed such as the Langmuir [25], the Szyszkowski [26] and the Frumkin [27] equations. The Gibbs equation itself has been simplified by Guggenheim and Adam with the choice of a different dividing plane and where the interfacial region is considered as a separate bulk phase (of finite volume) [28].

### 4.3.3 Efficiency and effectiveness of surfactant adsorption

The performance of a surfactant in lowering the surface tension of a solution can be discussed in terms of (1) the concentration required to produce a given surface tension reduction and (2) the maximum reduction in surface tension that can be obtained regardless of the concentration. These are referred to as the surfactant *efficiency* and *effectiveness* respectively.

A good measure of the *surfactant adsorption efficiency* is the concentration of surfactant required to produce a  $20 \text{ mN m}^{-1}$  reduction in surface tension. At this value the surfactant concentration is close to the minimum concentration needed to produce maximum adsorption at the interface. This is confirmed by the Frumkin adsorption equation (4.16), which relates the reduction in surface tension (or surface pressure  $\pi$ ) and surface excess concentration:

$$\gamma_o - \gamma = \pi = -2.303 RT \Gamma_m \log \left( 1 - \frac{\Gamma_1}{\Gamma_m} \right) \quad (4.16)$$

The maximum surface excess generally lies in the range  $1\text{--}4.4 \times 10^{-10} \text{ mol cm}^{-2}$  [29]: solving Equation (4.16) indicates that when the surface tension has been reduced by  $20 \text{ mN m}^{-1}$ , at  $25^\circ\text{C}$ , the surface is 84–99.9% saturated. The negative logarithm of such concentration,  $\text{p}C_{20}$ , is then a useful quantity since it can be related to the free energy change  $\Delta G^\circ$  involved in the transfer of a surfactant molecule from the interior of the bulk liquid phase to the interface. The surfactant adsorption efficiency thus relates to the structural groups in the molecule via the standard free energy change of the individual groups (i.e. free energies of transfer of methylene, terminal methyl and head groups). In particular, for a given homologous series of straight-chain surfactants in water,  $\text{CH}_3(\text{CH}_2)_n\text{--M}$ , where M is the hydrophilic head group and  $n$  is the number of methylene units in the chain, and when the systems are at  $\pi = 20 \text{ mN m}^{-1}$ , the standard free energy of adsorption is

$$\Delta G^\circ = n\Delta G^\circ(-\text{CH}_2-) + \Delta G^\circ(\text{M}) + \Delta G^\circ(\text{CH}_3-) \quad (4.17)$$

Then the adsorption efficiency is directly related to the length of the hydrophobic chain (the hydrophilic group remains the same), viz.

$$-\log(C)_{20} = \text{p}C_{20} = n \left[ \frac{-\Delta G^\circ(-\text{CH}_2-)}{2.303 RT} \right] + \text{const} \quad (4.18)$$

$\Delta G^\circ(\text{M})$  is considered as a constant and it is assumed that  $\Gamma_m$  does not differ significantly with increasing chain length, and that activity coefficients are unity. The efficiency factor

$pC_{20}$  therefore increases linearly with the number of carbon atoms in the hydrophobic chain. This is also described by Traube's rule [30] (Equation (4.19)):

$$\text{Log } C_s = B - n \text{Log } K_T \quad (4.19)$$

$C_s$  is the surfactant concentration,  $B$  is a constant,  $n$  is the chain length within a homologous series and  $K_T$  is Traube's constant. For hydrocarbon straight chain surfactants  $K_T$  is usually around 3 [31] or by analogy to Equation (4.18) is given by

$$\frac{C_n}{C_{n+1}} = K_T = \exp \left[ \frac{-\Delta G^\circ(-\text{CH}_2-)}{2RT} \right] \quad (4.20)$$

For compounds having a phenyl group in the hydrophobic chain it is equivalent to about three and one-half normal  $-\text{CH}_2-$  groups.

The larger the  $pC_{20}$  the more efficiently the surfactant adsorbed at the interface and the more efficiently it reduces surface tension. The other main factors that contribute to an increase in surfactant efficiency are summarised below:

- A straight alkyl chain as the hydrophobic group, rather than a branched alkyl chain containing the same number of carbon atoms
- A single hydrophilic group situated at the end of the hydrophobic group, rather than one (or more) at a central position
- A non-ionic or zwitterionic hydrophilic group, rather than an ionic one
- For ionic surfactants, a reduction in the effective charge by (a) use of a more tightly bound (less hydrated) counterion and (b) increase in ionic strength of the aqueous phase.

The choice of  $20 \text{ mN m}^{-1}$  as a standard value of surface tension lowering for the definition of adsorption efficiency is convenient but somewhat arbitrary, and is not valid for systems where surfactants differ significantly in maximum surface excess or when the surface pressure is less than  $20 \text{ mN m}^{-1}$ . Pitt *et al.* [32] circumvented this problem by defining  $\Delta\gamma$  as half the surface pressure at the CMC.

The performance of a surfactant can also be discussed in terms of *effectiveness of adsorption*. This is usually defined as the maximum lowering of surface tension  $\gamma_{\min}$  regardless of concentration, or as the surface excess concentration at surface saturation  $\Gamma_m$  since it represents the maximum adsorption.  $\gamma_{\min}$  and  $\Gamma_m$  are controlled mainly by the critical micelle concentration, and for certain ionics by the solubility limit or Krafft temperature  $T_k$ , which will be described briefly in Section 4.4.1. The effectiveness of adsorption is an important factor in determining such properties as foaming, wetting and emulsification, since  $\Gamma_m$  through the Gibbs adsorption equation gives a measure of the interfacial packing.

The efficiency and effectiveness of surfactants do not necessarily run parallel, and it is commonly observed – as shown by Rosen's extensive data listing [29] – that materials producing significant lowering of the surface tension at low concentrations (i.e. they are more efficient) have smaller  $\Gamma_m$  (i.e. they are less effective). In determining surfactant efficiency the role of the molecular structure is primarily thermodynamic, while its role in effectiveness is directly related to the relative size of the hydrophilic and hydrophobic portions of the adsorbing molecule. The area occupied by each molecule is determined either by the hydrophobic chain cross-sectional area, or the area required for closest packing of head groups, whichever is greater. Therefore, surfactant films can be tightly or loosely packed resulting in very different interfacial properties. For instance, straight chains and large head

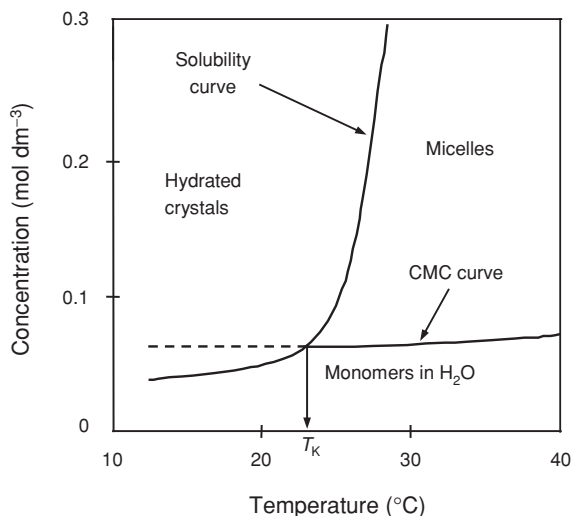
groups (relative to the tail cross section) favour close, effective packing, while branched, bulky or multiple hydrophobic chains give rise to steric hindrance at the interface. On the other hand, within a series of single straight chain surfactants, increasing the hydrocarbon chain length from C8 to C20 will have little effect on adsorption effectiveness [29].

## 4.4 Surfactant Solubility

In aqueous solution, when all available interfaces are saturated, the overall energy reduction may continue through other mechanisms. Depending on the system composition, a surfactant molecule can play different roles in terms of aggregation (formation of micelles, liquid crystal phases, bilayers or vesicles, etc.). The physical manifestation of one such mechanism is crystallisation or precipitation of surfactant from solution – that is, bulk-phase separation. While most common surfactants have a substantial solubility in water, this can change significantly with variations in hydrophobic tail length, head group nature, counterion valence, solution environment and most importantly, temperature.

### 4.4.1 The Krafft temperature

As for most solutes in water, increasing temperature produces an increase in solubility. However, for ionic surfactants, which are initially insoluble, there is often a temperature at which the solubility suddenly increases very dramatically. This is known as the Krafft point or Krafft temperature,  $T_K$ , and is defined as the intersection of the solubility and the CMC curves, i.e. it is the temperature at which the solubility of the monomeric surfactant is equivalent to its CMC at the same temperature. This is illustrated in Figure 4.4. Below



**Figure 4.4** The Krafft temperature  $T_K$  is the point at which surfactant solubility equals the critical micelle concentration. Above  $T_K$ , surfactant molecules form a dispersed phase; below  $T_K$ , hydrated crystals are formed.

$T_K$ , surfactant monomers only exist in equilibrium with the hydrated crystalline phase, and above  $T_K$ , micelles are formed providing much greater surfactant solubility.

The Krafft point of ionic surfactants is found to vary with counterion [33], alkyl chain length and chain structure. Knowledge of the Krafft temperature is crucial in many applications since below  $T_K$  the surfactant will clearly not perform efficiently; hence typical characteristics such as maximum surface tension lowering and micelle formation cannot be achieved. The development of surfactants with a lower Krafft point but still being very efficient at lowering surface tension (i.e. long chain compounds) is usually achieved by introducing chain branching, multiple bonds in the alkyl chain or bulkier hydrophilic groups thereby reducing intermolecular interactions that would tend to promote crystallisation.

#### 4.4.2 The cloud point

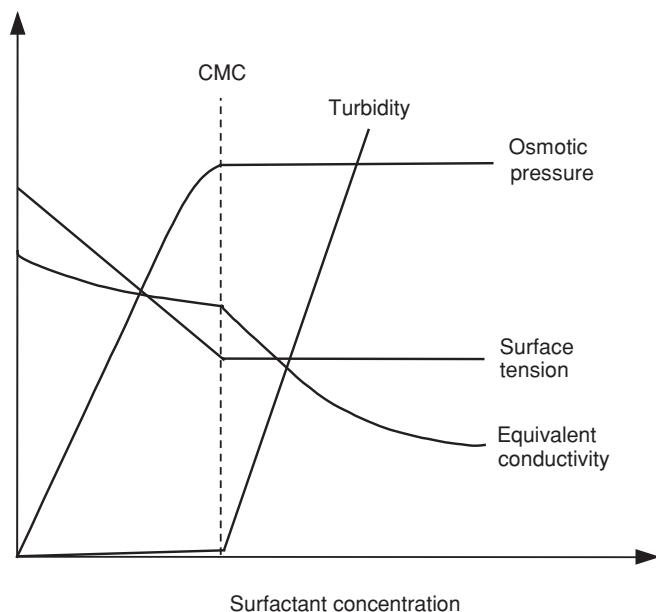
For non-ionic surfactants, a common observation is that micellar solutions tend to become visibly turbid at a well-defined temperature. This is often referred to as the cloud point, above which the surfactant solution phase separates. Above the cloud point, the system consists of an almost micelle-free dilute solution at a concentration equal to its CMC at that temperature, and a surfactant-rich micellar phase. This separation is caused by a sharp increase in aggregation number and a decrease in intermicellar repulsions [34, 35] that produces a difference in density of the micelle-rich and micelle-poor phases. Since much larger particles are formed, the solution becomes visibly turbid with large micelles efficiently scattering light. As with Krafft temperatures, the cloud point depends on chemical structure. For polyoxyethylene (PEO) non-ionics, the cloud point increases with increasing EO content for a given hydrophobic group, and at constant EO content it may be lowered by decreasing the hydrophobe size, broadening the PEO chain-length distribution and branching in the hydrophobic group [36].

### 4.5 Micellisation

In addition to forming oriented interfacial monolayers, surfactants can aggregate to form *micelles*, provided their concentration is sufficiently high. They are typically clusters of 50–200 surfactant molecules, whose size and shape are governed by geometric and energetic considerations. Micelle formation occurs over a fairly sharply defined region called the *critical micelle concentration*. Above the CMC, additional surfactant forms the aggregates, whereas the concentration of the unassociated monomers remains almost constant. As a result, a rather abrupt change in concentration dependence at much the same point can be observed in common equilibrium or transport properties (Figure 4.5).

#### 4.5.1 Thermodynamics of micellisation

Micelles are dynamic species, in that there is a constant, rapid interchange – typically on a microsecond timescale – of molecules between the aggregate and solution pseudo-phases.



**Figure 4.5** Schematic representation of the concentration dependence of some physical properties for solutions of a micelle-forming surfactant.

This constant formation-dissociation process relies on a subtle balance of interactions. These come from contacts between (1) hydrocarbon chain and water, (2) hydrocarbon and hydrocarbon chains, (3) head group and head group and (4) from solvation of the head group. Therefore, the net free energy change upon micellisation,  $\Delta G_m$ , can be written as

$$\Delta G_m = \Delta G(\text{HC}) + \Delta G(\text{contact}) + \Delta G(\text{packing}) + \Delta G(\text{HG}) \quad (4.21)$$

where

- $\Delta G(\text{HC})$  is the free energy associated with transferring hydrocarbon chains out of water and into the oil-like interior of the micelle.
- $\Delta G(\text{contact})$  is a surface free energy attributed to solvent–hydrocarbon contacts in the micelle.
- $\Delta G(\text{packing})$  is a positive contribution associated with confining the hydrocarbon chain to the micelle core.
- $\Delta G(\text{HG})$  is a positive contribution associated with head group interactions, including electrostatic as well as head group conformation effects.

Aggregation of surfactant molecules partly results from the tendency of the hydrophobic groups to minimise contacts with water by forming oily microdomains within the solvent. There, alkyl–alkyl interactions are maximised, while hydrophilic head groups remain surrounded by water.

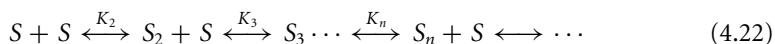
The traditional picture of micelle formation thermodynamics is based on the Gibbs–Helmholtz equation ( $\Delta G_m = \Delta H_m - T\Delta S_m$ ). At room temperature the process is characterised by a small, positive enthalpy,  $\Delta H_m$ , and a large, positive entropy of micellisation,

$\Delta S_m$ . The latter is considered as the main contribution to the negative  $\Delta G_m$  value, and so has led to the controversial idea that micellisation is an entropy-driven process. High positive values of  $\Delta S_m$  are indeed surprising since aggregation, in terms of configurational entropy, should result in a negative contribution (i.e. formation of ordered aggregates from free surfactant monomers). In addition, large values of  $\Delta H_m$  would have been expected since hydrocarbon groups have very little solubility in water, and consequently a high enthalpy of solution.

One mechanism that accounts for such conflicts is that when alkyl groups are surrounded by water, the  $H_2O$  molecules form clathrate cavities (i.e. stoichiometric crystalline solids in which water forms cages around solutes), thereby increasing either the strength or number of effective hydrogen bonds [37]. Therefore, the predominant effect of the hydrocarbon molecule is to increase the degree of structure in the immediately surrounding water. This is one of the main features of the *hydrophobic effect*, a subject that was explored in detail by Tanford [2] to account for the very slight solubility of hydrocarbons in water. During the formation of micelles, the reverse process occurs: as lyophobic residues aggregate, the highly structured water around each chain collapses back to ordinary bulk water thereby accounting for the apparent large overall gain in entropy,  $\Delta S_m$ . This water-structure effect was also invoked by other researchers [38, 39].

Such an interpretation, however, has been strongly challenged by more recent studies of aqueous systems at high temperatures (up to 166°C) and micellisation in hydrazine solutions [40]. In these systems water loses most of its peculiar structural properties and the formation of structured water around lyophobic species is no longer possible.

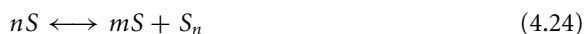
The mechanism of micelle formation from surfactant monomers,  $S$ , can be described by a series of step-wise equilibria:



with equilibrium constants  $K_n$  for  $n = 2 - \infty$ , and where the various thermodynamic parameters ( $\Delta G^\circ$ ,  $\Delta H^\circ$ ,  $\Delta S^\circ$ ) for the aggregation process can be expressed in terms of  $K_n$ . However, each  $K_n$  cannot be measured individually, so different approaches have been proposed to model the energetics of the process of self-association. Although not totally accurate, two simple models are generally encountered: the closed-association and the phase separation models. In the closed-association model, with the size range of spherical micelles around the CMC being very limited, it is assumed that only one of  $K_n$  value is dominant, and micelles and monomeric species are considered to be in chemical equilibrium:



$n$  is the number of molecules of surfactant,  $S$ , associating to form the micelle (i.e. the aggregation number). In the phase separation model, the micelles are considered to form a new phase within the system at and above the critical micelle concentration, and



where  $m$  is the number of free surfactant molecules in the solution and  $S_n$  the new phase. In both cases, equilibrium between monomeric surfactant and micelles is assumed with a

corresponding equilibrium constant,  $K_m$ , given by

$$K_m = \frac{[\text{micelles}]}{[\text{monomers}]^n} = \frac{[S_n]}{[S]^n} \quad (4.25)$$

where brackets indicate molar concentrations and  $n$  is the number of monomers in the micelle, the aggregation number. Although micellisation is itself a source of non-ideality [41, 42], it is assumed in Equation (4.25) that activities may be replaced by concentrations.

From Equation (4.25), the standard free energy of micellisation per mole of micelles is given by

$$\Delta G_m^\circ = -RT \ln K_m = -RT \ln S_n + nRT \ln S \quad (4.26)$$

while the standard free energy change per mole of surfactant is

$$\frac{\Delta G_m^\circ}{n} = -\frac{RT}{n} \ln S_n + RT \ln S \quad (4.27)$$

Assuming  $n$  is large ( $\sim 100$ ) the first term on the right side of Equation (4.27) can be neglected, and an approximate expression for the free energy of micellisation per mole of a neutral surfactant becomes

$$\Delta G_{M,m}^\circ \approx RT \ln(\text{CMC}) \quad (4.28)$$

In the case of ionic surfactants, the presence of the counterion and its degree of association with the monomer and micelle must be considered. The mass-action equation becomes

$$nS^x + (n-p)C^y \leftrightarrow S_n^\alpha \quad (4.29)$$

where  $C$  is the concentration of free counterions. The degree of dissociation of the surfactant molecules in the micelle,  $\alpha$ , the micellar charge, is given by  $\alpha = p/n$ .

The ionic equivalent to Equation (4.25) is then

$$K_m = \frac{[S_n]}{[S^x]^n [C^y]^{(n-p)}} \quad (4.30)$$

where  $p$  is the concentration of free counterions associated with, but not bound to the micelle. The standard free energy of micelle formation becomes

$$\Delta G_m^\circ = -RT \{ \ln[S_n] - n \ln[S^x] - (n-p) \ln[C^y] \} \quad (4.31)$$

At the CMC  $[S^{(-)}] = [C^{(+)}] = \text{CMC}$  for a fully ionised surfactant, and the standard free energy change per mole of surfactant can be obtained from the approximation:

$$\Delta G_{M,m}^\circ \approx RT \left( 2 - \frac{p}{n} \right) \ln(\text{CMC}) \quad (4.32)$$

When the ionic micelle is in a solution of high electrolyte content, the situation described by Equation (4.32) reverts to the simple non-ionic case given by Equation (4.28).

From the Gibbs function and second law of thermodynamics,  $\Delta S^\circ$  for non-ionic surfactants is given as

$$\Delta S^\circ = -\frac{d(\Delta G^\circ)}{dT} = -RT \frac{d \ln(\text{CMC})}{dT} - R \ln(\text{CMC}) \quad (4.33)$$



From the Gibbs function and Equations (4.28) and (4.33), the enthalpy of micellisation for non-ionic surfactants,  $\Delta H^\circ$ , is given by

$$\Delta H^\circ = \Delta G^\circ + T\Delta S^\circ = -RT^2 \frac{d\ln(\text{CMC})}{dT} \quad (4.34)$$

and similarly for ionics,

$$\Delta H^\circ = -RT^2 \left(2 - \frac{p}{n}\right) \frac{d\ln(\text{CMC})}{dT} \quad (4.35)$$

Both the phase separation and closed association models have disadvantages. One difficulty is activity coefficients: assuming ideality can be erroneous considering the large effective micelle size and charge in comparison to dilute solutions of surfactant monomers. However, the model described above is useful enough to be applied to the systems presented in this study. Another disadvantage is the assumption of micellar monodispersity. To counteract this problem, the multiple equilibrium model was proposed, which is an extension of the closed association model. It allows a distribution function of aggregation numbers in micelles to be calculated. A full account of this model and its derivation can be found in [43–45].

#### 4.5.2 Factors affecting the CMC

Many factors are known to strongly affect the CMC. Of major effect is the structure of the surfactant, as will be described below. Also important, but to a lesser extent, are parameters such as counterion nature, presence of additives and change in temperature.

*The hydrophobic group: the ‘tail’.* The length of the hydrocarbon chain is a major factor determining the CMC. For a homologous series of linear single-chain surfactants the CMC decreases logarithmically with carbon number. The relationship usually fits the Klevens equation [46]

$$\log_{10}(\text{CMC}) = A - Bn_c \quad (4.36)$$

where  $A$  and  $B$  are constants for a particular homologous series and temperature, and  $n_c$  is the number of carbon atoms in the chain,  $C_nH_{2n+1}$ . The constant  $A$  varies with the nature and number of hydrophilic groups, while  $B$  is constant and approximately equal to  $\log_{10}2$  ( $B \approx 0.29$ – $0.30$ ) for all paraffin chain salts having a single ionic head group (i.e. reducing the CMC to approximately one-half per each additional  $-\text{CH}_2-$  group).

Interestingly, for straight-chain dialkyl sulfosuccinates Equation (4.36) is still valid [47] and  $B \approx 0.62$ , which essentially doubles the value for the single chain compounds. Alkyl chain branching and double bonds, aromatic groups or some other polar character in the hydrophobic part produces noticeable changes in CMC. In hydrocarbon surfactants, chain branching gives a higher CMC than a comparable straight chain surfactant [29], and introduction of a benzene ring in the chain is equivalent to about 3.5 carbon atoms.

*The hydrophilic group.* For surfactants with the same hydrocarbon chain, varying the hydrophile nature (i.e. from ionic to non-ionic) has an important effect on the CMC values. For instance, for a C12 hydrocarbon the CMC with an ionic head group lies in the range of  $1 \times 10^{-3} \text{ mol dm}^{-3}$ , while a C12 non-ionic material exhibits a CMC in the range of  $1 \times 10^{-4} \text{ mol dm}^{-3}$ . The exact nature of the ionic group, however, has no dramatic effect, since a major driving force for micelle formation is the entropy factor discussed above.

*Counterion effects.* In ionic surfactants micelle formation is related to the interactions of solvent with the ionic head group. Since electrostatic repulsions between ionic groups are greatest for complete ionisation, an increase in the degree of ion binding will decrease the CMC. For a given hydrophobic tail and anionic head group, the CMC decreases as  $\text{Li}^+ > \text{Na}^+ > \text{K}^+ > \text{Cs}^+ > \text{N}(\text{CH}_3)_4^+ > \text{N}(\text{CH}_2\text{CH}_3)_4^+ > \text{Ca}^{2+} \approx \text{Mg}^{2+}$ . For cationic series such as the dodecyltrimethylammonium halides, the CMC decreases in the order  $\text{F}^- > \text{Cl}^- > \text{Br}^- > \text{I}^-$ . In addition, varying counterion valency produces a significant effect. Changing from monovalent to di- or trivalent counterions produces a sharp decrease in the CMC.

*Effect of added salt.* The presence of an indifferent electrolyte causes a decrease in the CMC of most surfactants. The greatest effect is found for ionic materials. The principal effect of the salt is to partially screen the electrostatic repulsion between the head groups and so lower the CMC. For ionics, the effect of adding electrolyte can be empirically quantified viz.

$$\log_{10}(\text{CMC}) = -a \log_{10} C_i + b \quad (4.37)$$

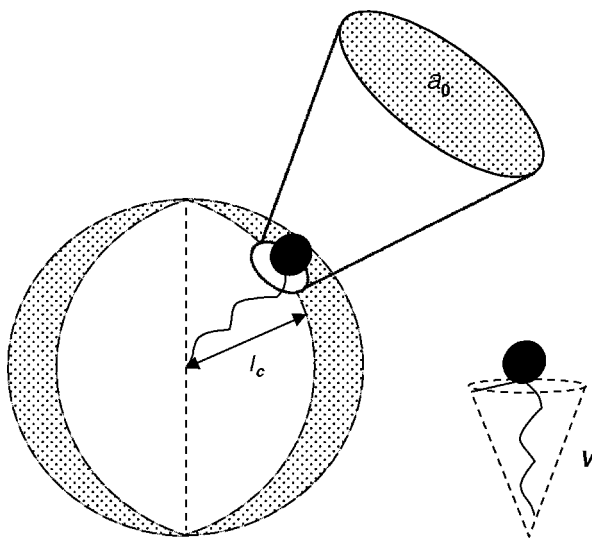
Non-ionic and zwitterionic surfactants display a much smaller effect and Equation (4.37) does not apply.

*Effect of temperature.* The influence of temperature on micellisation is usually weak, reflecting subtle changes in bonding, heat capacity and volume that accompany the transition. This is, however, quite a complex effect. It was shown, for example, that the CMC of most ionic surfactants passes through a minimum as the temperature is varied from 0 to 70°C [48]. As already mentioned (Section 4.4), the major effects of temperature are the Krafft and cloud points.

### 4.5.3 Structure of micelles and molecular packing

Early studies [49, 50] showed that with ionic single alkyl chain compounds spherical micelles form. In particular, in 1936 Hartley [51] described such micelles as spherical aggregates whose alkyl groups form a hydrocarbon liquid-like core, and whose polar groups form a charged surface. Later, with the development of zwitterionic and non-ionic surfactants, micelles of very different shapes were encountered. The different geometries were found to depend mainly on the structure of the surfactant, as well as environmental conditions (e.g. concentration, temperature, pH, electrolyte content).

In the micellisation process, molecular geometry plays an important role and it is essential to understand how surfactants can pack. The main structures encountered are spherical



**Figure 4.6** The critical packing parameter  $P_c$  (or surfactant number) relates the head group area, the extended length and the volume of the hydrophobic part of a surfactant molecule into a dimensionless number  $P_c = v/a_0 l_c$ .

micelles, vesicles, bilayers or inverted micelles. As described previously, two opposing forces control the self-association process: hydrocarbon–water interactions that favour aggregation (i.e. pulling surfactant molecules out of the aqueous environment), and head group interactions that work in the opposite sense. These two contributions can be considered as an attractive interfacial tension term due to hydrocarbon tails and a repulsion term depending on the nature of the hydrophilic group. More recently, this basic idea was reviewed and quantified by Mitchell and Ninham [52] and Israelachvili [53], resulting in the concept that aggregation of surfactants is controlled by a balanced molecular geometry. In brief, the geometric treatment separates the overall free energy of association to three critical geometric terms (Figure 4.6):

- the minimum interfacial area occupied by the head group,  $a_0$
- the volume of the hydrophobic tail(s),  $v$
- the maximum extended chain length of the tail in the micelle core,  $l_c$ .

Formation of a spherical micelle requires  $l_c$  to be equal to or less than the micelle core radius,  $R_{\text{mic}}$ . Then for such a shape, an aggregation number,  $N$ , can be expressed either as the ratio of micellar core volume,  $V_{\text{mic}}$ , and that for the tail,  $v$ :

$$N = V_{\text{mic}}/v = [(4/3)\pi R_{\text{mic}}^3]/v \quad (4.38)$$

or as the ratio between the micellar area,  $A_{\text{mic}}$ , and the cross-sectional area,  $a_0$ :

$$N = A_{\text{mic}}/a_0 = [4\pi R_{\text{mic}}^2]/a_0 \quad (4.39)$$

Equating Equations (4.38) and (4.39)

$$v/(a_0 R_{\text{mic}}) = 1/3 \quad (4.40)$$

**Table 4.4** Expected aggregate characteristics in relation to surfactant critical packing parameter,  $P_c = v/a_0l_c$ 

$P_c$	General Surfactant type	Expected Aggregate Structure
<0.33	Single-chain surfactants with large head groups	Spherical or ellipsoidal micelles
0.33–0.5	Single-chain surfactants with small head groups, or ionics in the presence of large amounts of electrolyte	Large cylindrical or rod-shaped micelles
0.5–1.0	Double-chain surfactants with large head groups and flexible chains	Vesicles and flexible bilayers structures
1.0	Double-chain surfactants with small head groups or rigid, immobile chains	Planar extended bilayers
>1.0	Double-chain surfactants with small head groups, very large and bulky hydrophobic groups	Reversed or inverted micelles

Since  $l_c$  cannot exceed  $R_{\text{mic}}$  for a spherical micelle

$$v/(a_0l_c) \leq 1/3 \quad (4.41)$$

More generally, this defines a critical packing parameter,  $P_c$ , as the ratio of volume to surface area:

$$P_c = v/(a_0l_c) \quad (4.42)$$

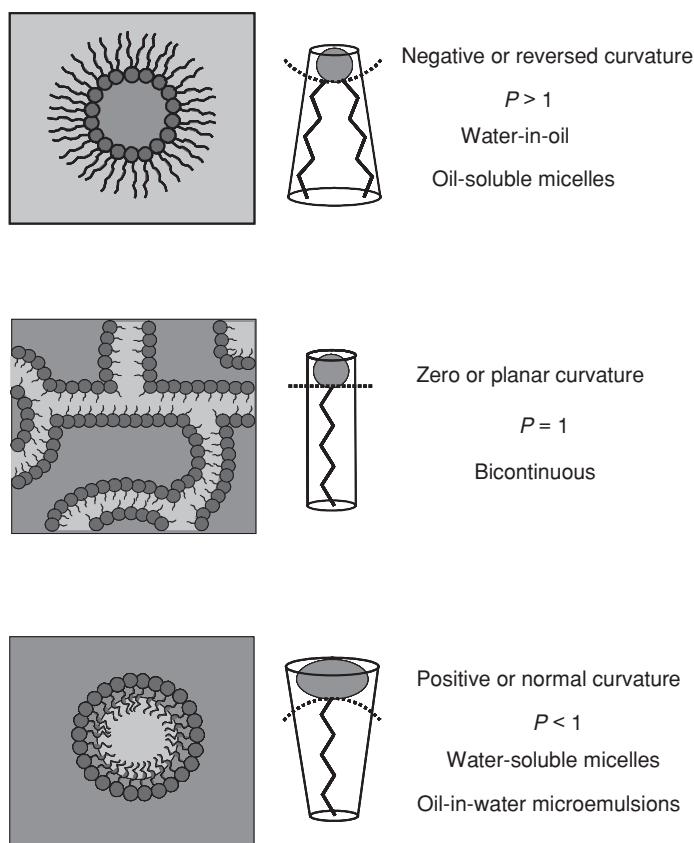
The parameter  $v$  varies with the number of hydrophobic groups, chain unsaturation, chain branching and chain penetration by other compatible hydrophobic groups, while  $a_0$  is mainly governed by electrostatic interactions and head group hydration.  $P_c$  is a useful quantity since it allows the prediction of aggregate shape and size. The predicted aggregation characteristics of surfactants cover a wide range of geometric possibilities, and the main types are presented in Table 4.4 and Figures 4.7 and 4.8.

## 4.6 Liquid Crystalline Mesophases

Micellar solutions, although the subject of extensive studies and theoretical considerations, are only one of several possible aggregation states. A complete understanding of the aqueous behaviour of surfactants requires knowledge of the entire spectrum of self-assembly. The existence of liquid crystalline phases constitutes an equally important aspect and a detailed description can be found in the literature [e.g. 54, 55]. The common features of liquid crystalline phases are summarised below.

### 4.6.1 Definition

When the volume fraction of surfactant in a micellar solution is increased, typically above a threshold of about 40%, a series of regular geometries is commonly encountered. Interactions between micellar surfaces are repulsive (from electrostatic or hydration forces), so

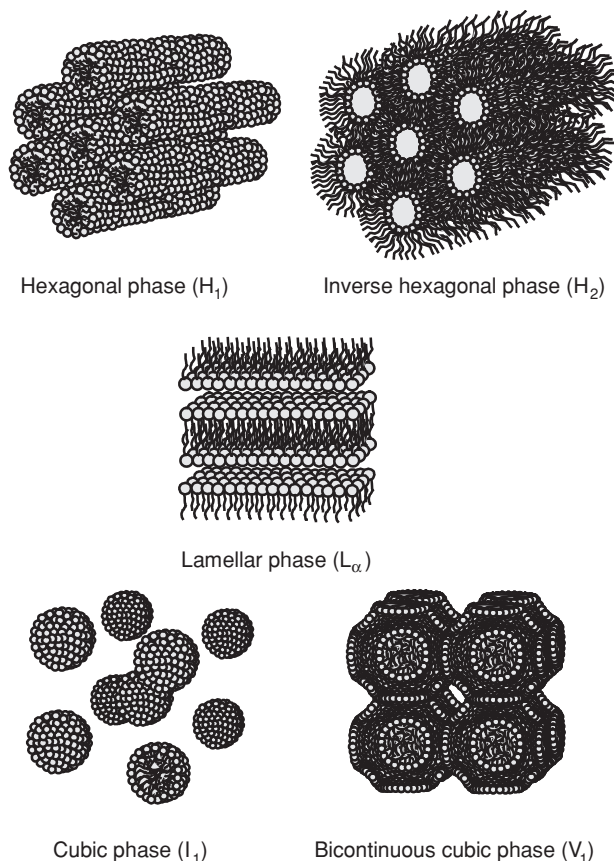


**Figure 4.7** Changes in the critical packing parameters ( $P_c$ ) of surfactant molecules give rise to different aggregation structures.

that as the number of aggregates increases and micelles get closer to one another, the only way to maximise separation is to change shape and size. This explains the sequence of surfactant phases observed in the concentrated regime. Such phases are known as mesophases or lyotropic (solvent-induced) liquid crystals.

As the term suggests, liquid crystals are characterised by having physical properties intermediate between crystalline and fluid structures: the degree of molecular ordering is between that of a liquid and a crystal and in terms of rheology the systems are neither simple viscous liquids nor crystalline elastic solids. Some of these phases have at least one direction that is highly ordered so that liquid crystals exhibit optical birefringence.

Two general classes are encountered depending on whether one is considering surfactants or other types of material. These are *thermotropic* liquid crystals, in which the structure and properties are determined by temperature (such as employed in LCD cells). For *lyotropic* liquid crystals structure is determined by specific interactions between solute and solvent: surfactant liquid crystals are normally lyotropic.



**Figure 4.8** Common surfactant liquid crystalline phases. See Table 1.4 for identification.

#### 4.6.2 Structures

The main structures associated with two-component surfactant–water systems are: hexagonal (normal or inverted), lamellar and several cubic phases. Table 4.5 summarises the notations commonly associated with these phases and their structures are shown in Figure 4.8.

- The *hexagonal phase* is composed of a close-packed array of long cylindrical micelles, arranged in a hexagonal pattern. The micelles may be ‘normal’ (in water,  $H_1$ ) in that the hydrophilic head groups are located on the outer surface of the cylinder, or ‘inverted’ ( $H_2$ ), with the hydrophilic group located internally. Since all the space between adjacent cylinders is filled with hydrophobic groups, the cylindrical micelles are more closely packed than those found in the  $H_1$  phase. As a result,  $H_2$  phases occupy a much smaller region of the phase diagram and are much less common.
- The *lamellar phase* ( $L_\alpha$ ) is built up of alternating water-surfactant bilayers. The hydrophobic chains possess a significant degree of randomness and mobility, and the surfactant

**Table 4.5** Most common lyotropic liquid crystalline and other phases found in binary surfactant–water systems

Phase structure	Symbol	Other names
Lamellar	$L_\alpha$	Neat
Hexagonal	$H_1$	Middle
Reversed hexagonal	$H_2$	
Cubic (normal micellar)	$I_1$	Viscous isotropic
Cubic (reversed micellar)	$I_2$	
Cubic (normal bicontinuous)	$V_1$	Viscous isotropic
Cubic (reversed bicontinuous)	$V_2$	
Micellar	$L_1$	
Reversed micellar	$L_2$	

bilayer can range from being stiff and planar to being very flexible and undulating. The level of disorder may vary smoothly or change abruptly, depending on the specific system, so that it is possible for a surfactant to pass through several distinct lamellar phases.

- The *cubic phase* may have a wide variety of structural variations and occurs in many different parts of the phase diagram. These are optically isotropic systems and so cannot be characterised by polarising light microscopy. Two main groups of cubic phases have been identified:
  - i. The micellar cubic phases ( $I_1$  and  $I_2$ ) built up of regular packing of small micelles (or reversed micelles in the case of  $I_2$ ). The micelles are short prolates arranged in a body-centred cubic close-packed array [56, 57].
  - ii. The bicontinuous cubic phases ( $V_1$  and  $V_2$ ) are thought to be rather extended, porous, connected structures in three dimensions. They are considered to be formed by either connected rod-like micelles, similar to branched micelles, or bilayer structures. Denoted  $V_1$  and  $V_2$ , they can be normal or reverse structures and are positioned between  $H_1$  and  $L_\alpha$  and between  $L_\alpha$  and  $H_2$  respectively.

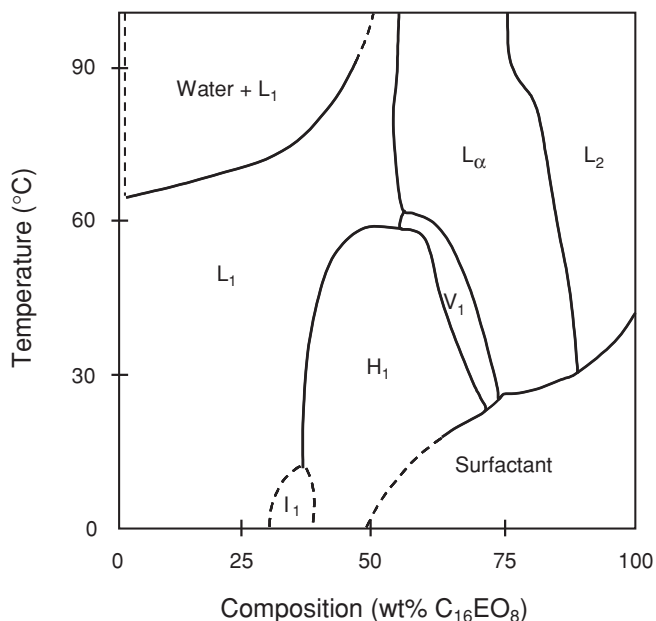
In addition to having different structures these common forms also show different viscosities, in the order

$$\text{Cubic} > \text{Hexagonal} > \text{Lamellar}$$

Cubic phases are generally the more viscous since they have no obvious shear plane and so layers of surfactant aggregates cannot slide easily relative to each other. Hexagonal phases typically contain 30–60% water by weight but are very viscous since cylindrical aggregates can move freely only along their length. Lamellar phases are generally less viscous than the hexagonal phases due to the ease with which each parallel layer can slide over each other during shear.

#### 4.6.3 Phase diagrams

The sequence of mesophases can be identified simply by using a polarising microscope and the isothermal technique known as a phase cut. Briefly, starting from a small amount of



**Figure 4.9** Phase diagram for the non-ionic  $C_{16}EO_8$  illustrating the various liquid crystalline phases.  $L_1$  and  $L_2$  are isotropic solutions. See Table 4.5 for details of the other phases. (After Mitchell D. J. *et al.* (1983) *J. Chem. Soc. Faraday Trans. I*, **79**, 975. Reproduced by permission of The Royal Society of Chemistry.)

surfactant, a concentration gradient is set up spanning the entire phase diagram, from pure water to pure surfactant. Since crystal hydrates and some of the liquid crystalline phases are birefringent, viewing in the microscope between crossed polars shows up the complete sequence of mesophases.

Transformations between different mesophases are controlled by a balance between molecular packing geometry and inter-aggregate forces. As a result, the system characteristics are highly dependent on the nature and amount of solvent present. Generally, the main types of mesophases tend to occur in the same order and in roughly the same position in the phase diagram. Figure 4.9 shows a classic binary phase diagram of a non-ionic surfactant  $C_{16}EO_8$ –water. The sequence of phases is common to most non-ionic surfactants of the kind  $C_iE_j$ , although the positions of the phase boundaries, in terms of temperature and concentration limits, depend somewhat on the chemical identity of the surfactant.

## References

1. Evans, D.F. and Wennerström, H. (1999) *The Colloidal Domain*. Wiley-VCH, New York.
2. Tanford, C. (1978) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*. Wiley, New York.
3. Ogino, K. and Abe, M. (eds) (1993) *Mixed Surfactants Systems*. Dekker, New York.
4. Robb, I.D. (1997) *Specialist Surfactants*. Blackie Academic, London.
5. Holmberg, K. (ed.) (1998) *Novel Surfactants*. Dekker, New York.



6. Hollis, G. (ed.) (1976) *Surfactants UK*. Tergo-Data.
7. The Soap and Detergent Association home page, <http://www.sdahq.org/>
8. Karsa, D.R., Goode, J.M. and Donnelly, P.J. (eds) (1991) *Surfactants Applications Directory*. Blackie, London.
9. Dickinson, E. (1992) *An Introduction to Food Colloids*. Oxford University Press, Oxford.
10. Solans, C. and Kunieda, H. (eds) (1997) *Industrial Applications of Microemulsions*. Dekker, New York.
11. Tanford, C. (1978) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*. Wiley, New York.
12. Dukhin, S.S., Kretschmar, G. and Miller, R. (1995) *Dynamics of Adsorption at Liquid Interfaces*. Elsevier, Amsterdam.
13. Rusanov, A.I. and Prokhorov, V.A. (1996) *Interfacial Tensiometry*. Elsevier, Amsterdam.
14. Chang, C.-H. and Franses, E.I. (1995) *Colloid Surf.*, **100**, 1.
15. Miller, R., Joos, P. and Fainermann, V. (1994) *Adv. Colloid Interface Sci.*, **49**, 249.
16. Lin, S.-Y., McKeigue, K. and Maldarelli, C. (1991) *Langmuir*, **7**, 1055.
17. Hsu, C.-H., Chang, C.-H. and Lin, S.-Y. (1999) *Langmuir*, **15**, 1952.
18. Eastoe, J. and Dalton, J.S. (2000) *Adv. Colloid Interface Sci.*, **85**, 103.
19. Gibbs, J.W. (1931) *The Collected Works of J. W. Gibbs*. Longmans, Green, New York. vol. I, p. 219.
20. Elworthy, P.H. and Mysels, K.J. (1966) *J Colloid Interface Sci.*, **21**, 331.
21. Lu, J.R., Li, Z.X., Su, T.J., Thomas, R.K. and Penfold, J. (1993) *Langmuir*, **9**, 2408.
22. Bae, S., Haage, K., Wantke, K. and Motschmann, H. (1999) *J. Phys. Chem. B*, **103**, 1045.
23. Downer, A., Eastoe, J., Pitt, A.R., Penfold, J. and Heenan, R.K. (1999) *Colloids Surf. A*, **156**, 33.
24. Eastoe, J., Nave, S., Downer, A., Paul, A., Rankin, A., Tribe, K. and Penfold, J. (2000) *Langmuir*, **16**, 4511.
25. Langmuir, I. (1948) *J. Am. Chem. Soc.*, **39**, 1917.
26. Szyszkowski, B. (1908) *Z. Phys. Chem.*, **64**, 385.
27. Frumkin, A. (1925) *Z. Phys. Chem.*, **116**, 466.
28. Guggenheim, E.A. and Adam, N.K. (1933) *Proc. R. Soc. (London) A*, **139**, 218.
29. Rosen, M.J. (1989) *Surfactants and Interfacial Phenomena*. Wiley, New York.
30. Traube, I. (1891) *Justus Liebigs Ann. Chem.*, **265**, 27.
31. Tamaki, K., Yanagushi, T. and Hori, R. (1961) *Bull. Chem. Soc. Japan*, **34**, 237.
32. Pitt, A.R., Morley, S.D., Burbidge, N.J. and Quickenden, E.L. (1996) *Colloids Surf. A*, **114**, 321.
33. Hato, M., Tahara, M. and Suda, Y. (1979) *J. Colloids Interface Sci.*, **72**, 458.
34. Staples, E.J. and Tiddy, G.J.T. (1978) *J. Chem. Soc., Faraday Trans. 1*, **74**, 2530.
35. Tiddy, G.J.T. (1980) *Phys. Rep.*, **57**, 1.
36. Schott, H. (1969) *J. Pharm. Sci.*, **58**, 1443.
37. Frank, H.S. and Evans, M.W. (1945) *J. Chem. Phys.*, **13**, 507.
38. Evans, D.F. and Wightman, P.J. (1982) *J. Colloid Interface Sci.*, **86**, 515.
39. Patterson, D. and Barbe, M. (1976) *J. Phys. Chem.*, **80**, 2435.
40. Evans, D.F. (1988) *Langmuir*, **4**, 3.
41. Hunter, R.J. (1987) *Foundations of Colloid Science*. Oxford University Press, New York. vol. I.
42. Evans, D.F. and Ninham, B.W. (1986) *J. Phys. Chem.*, **90**, 226.
43. Corkhill, J.M., Goodman, J.F., Walker, T. and Wyer, J. (1969) *Proc. R. Soc. (London) A*, **312**, 243.
44. Mukerjee, P. (1972) *J. Phys. Chem.*, **76**, 565.
45. Aniansson, E.A.G. and Wall, S.N. (1974) *J. Phys. Chem.*, **78**, 1024.
46. Klevens, H. (1953) *J. Am. Oil Chem. Soc.*, **30**, 7, 4.
47. Williams, E.F., Woodberry, N.T. and Dixon, J.K. (1957) *J. Colloid Interface Sci.*, **12**, 452.
48. Krescheck, G.C. (1975) In F. Franks (ed.), *Water-A Comprehensive Treatise*. Plenum, New York. pp. 95–167.
49. McBain, J.W. (1913) *Trans. Faraday Soc.*, **9**, 99.

50. Reyhler, A. (1913) *Kolloid-Z.*, **12**, 283.
51. Hartley, G.S. (1936) *Aqueous Solutions of Paraffin Chain Salts*. Hermann & Cie, Paris.
52. Mitchell, D.J. and Ninham, B.W. (1981) *J. Chem. Soc. Faraday Trans. 2*, **77**, 601.
53. Israelachvili, J.N. (1985) *Intermolecular and Surface Forces*. Academic, London. p. 251.
54. Laughlin, R.G. (1994) *The Aqueous Phase Behaviour of Surfactants*. Academic, London.
55. Chandrasekhar, S. (1992) *Liquid Crystals*. Cambridge University Press, New York.
56. Fontell, K., Kox, K.K. and Hansson, E. (1985) *Mol. Cryst. Liquid Cryst. Lett.*, **1**, 9.
57. Fontell, K. (1990) *Colloid Polym. Sci.*, **268**, 264.

# Chapter 5

## Microemulsions

*Julian Eastoe*

This chapter is devoted to another important property of surfactants, that of stabilization of water–oil films and formation of microemulsions. These are a special kind of colloidal dispersion that have attracted a great deal of attention because of their ability to solubilise otherwise insoluble materials. Industrial applications of microemulsions have escalated in the last 40 years following an increased understanding of formation, stability and the role of surfactant molecular architecture.

### 5.1 Microemulsions: Definition and History

One of the best definitions of microemulsions is from Danielsson and Lindman [1] ‘a microemulsion is a system of water, oil and an amphiphile which is a single optically isotropic and thermodynamically stable liquid solution’. In some respects, microemulsions can be considered as small-scale versions of emulsions, i.e. droplet type dispersions either of oil-in-water (o/w) or of water-in-oil (w/o), with a size range of the order of 5–50 nm in drop radius. Such a description, however, lacks precision since there are significant differences between microemulsions and ordinary emulsions (or macroemulsions). In particular, in emulsions the average drop size grows continuously with time so that phase separation ultimately occurs under gravity force, i.e. they are thermodynamically unstable and their formation requires input of work. The drops of the dispersed phase are generally large ( $>0.1\ \mu\text{m}$ ) so that they often take on a milky, rather than a translucent appearance. On the other hand, microemulsions form spontaneously. As for simple aqueous systems, microemulsion formation is dependent on surfactant type and structure. If the surfactant is ionic and contains a single hydrocarbon chain (e.g. sodium dodecylsulfate, SDS) microemulsions are only formed if a co-surfactant (e.g. a medium size aliphatic alcohol) and/or electrolyte (e.g. 0.2 M NaCl) are also present. With double chain ionics (e.g. Aerosol-OT) and some non-ionic surfactants a co-surfactant is not necessary. This results from one of the most fundamental properties of microemulsions, that is, an ultra-low interfacial tension between the oil and water phases,  $\gamma_{\text{o/w}}$ . The main role of the surfactant is to reduce  $\gamma_{\text{o/w}}$  sufficiently – i.e. lowering the energy required to increase the interfacial area – so that spontaneous dispersion of water or oil droplets occurs and the system is thermodynamically stable. As described in Section 5.2.1 ultra-low tensions are crucial for the formation of microemulsions and depend on system composition.

Microemulsions were not really recognized until the work of Hoar and Schulman in 1943, who reported a spontaneous emulsion of water and oil on addition of a strong surface-active agent [2]. The term ‘microemulsion’ was first used even later by Schulman *et al.* [3] in 1959 to describe a multi-phase system consisting of water, oil, surfactant and alcohol, which forms a transparent solution. There has been much debate about the word ‘microemulsion’ to describe such systems [4]. Although not systematically used today, some prefer the names ‘nanoemulsion’, ‘micellar emulsion’ [5] or ‘swollen micelles’ [6]. Microemulsions were probably discovered well before the studies of Schulmann: Australian housewives have used since the beginning of last century water/eucalyptus oil/soap flake/white spirit mixtures to wash wool, and the first commercial microemulsions were probably the liquid waxes discovered by Rodawald in 1928. Interest in microemulsions really stepped up in the late 1970s and early 1980s when it was recognised that such systems could improve oil recovery and when oil prices reached levels where tertiary recovery methods became profit earning [7]. Nowadays this is no longer the case, but other microemulsion applications were discovered, e.g. catalysis, preparation of submicron particles, solar energy conversion and liquid–liquid extraction (mineral, proteins, etc.). Together with classical applications in detergency and lubrication, the field remains sufficiently important to continue to attract a number of scientists. From the fundamental research point of view, a great deal of progress has been made in the last 20 years in understanding microemulsion properties. In particular, interfacial film stability and microemulsion structures can now be characterised in detail owing to the development of new and powerful techniques such as small-angle neutron scattering (SANS, as described in Chapter 12). The following sections deal with fundamental microemulsion properties, i.e. formation and stability, surfactant films, classification and phase behaviour.

## 5.2 Theory of Formation and Stability

### 5.2.1 Interfacial tension in microemulsions

A simple picture for describing microemulsion formation is to consider a subdivision of the dispersed phase into very small droplets. Then the configurational entropy change,  $\Delta S_{\text{conf}}$ , can be approximately expressed as [8]

$$\Delta S_{\text{conf}} = -nk_{\text{B}}[\ln \phi + \{(1 - \phi)/\phi\} \ln(1 - \phi)] \quad (5.1)$$

where  $n$  is the number of droplets of dispersed phase,  $k_{\text{B}}$  is the Boltzmann constant and  $\phi$  is the dispersed phase volume fraction. The associated free energy change can be expressed as a sum of the free energy for creating new area of interface,  $\Delta A\gamma_{12}$ , and configurational entropy in the form [9]:

$$\Delta G_{\text{form}} = \Delta A\gamma_{12} - T\Delta S_{\text{conf}} \quad (5.2)$$

where  $\Delta A$  is the change in interfacial area  $A$  (equal to  $4\pi r^2$  per spherical droplet of radius  $r$ ) and  $\gamma_{12}$  is the interfacial tension between phases 1 and 2 (e.g. oil and water) at temperature  $T$  (K). Substituting Equation (5.1) into (5.2) gives an expression for obtaining the maximum interfacial tension between phases 1 and 2. On dispersion, the droplet number increases and  $\Delta S_{\text{conf}}$  is positive. If the surfactant can reduce the interfacial tension to a sufficiently low value, the energy term in Equation (5.2) ( $\Delta A\gamma_{12}$ ) will be relatively small and positive,

thus allowing a negative (and hence favourable) free energy change, that is, spontaneous microemulsification.

In surfactant-free oil–water systems,  $\gamma_{o/w}$  is of the order of  $50 \text{ mN m}^{-1}$ , and during microemulsion formation the increase in interfacial area,  $\Delta A$ , is very large, typically a factor of  $10^4$ – $10^5$ . Therefore in the absence of surfactant, the second term in Equation (5.2) is of the order of  $1000 k_B T$ , and in order to fulfil the condition  $\Delta A \gamma_{12} \leq T \Delta S_{\text{conf}}$ , the interfacial tension should be very low (approximately  $0.01 \text{ mN m}^{-1}$ ). Some surfactants (double chain ionics [10, 11] and some non-ionics [12]) can produce extremely low interfacial tensions – typically  $10^{-2} - 10^{-4} \text{ mN m}^{-1}$  – but in most cases, such low values cannot be achieved by a single surfactant. An effective way to further decrease  $\gamma_{o/w}$  is to include a second surface-active species (either a surfactant or medium-chain alcohol), that is a co-surfactant. This can be understood in terms of the Gibbs equation extended to multi-component systems [13–15]. It relates the interfacial tension to the surfactant film composition and the chemical potential,  $\mu$ , of each component in the system, i.e.

$$d\gamma_{o/w} = - \sum_i (\Gamma_i d\mu_i) \approx - \sum_i (\Gamma_i RT d \ln C_i) \quad (5.3)$$

where  $C_i$  is the concentration of component  $i$  in the mixture, and  $\Gamma_i$  the surface excess. Assuming that surfactants and co-surfactants, with concentrations  $C_s$  and  $C_{co}$  respectively, are the only adsorbed components (i.e.  $\Gamma_{\text{water}} = \Gamma_{\text{oil}} = 0$ ), Equation (5.3) becomes

$$d\gamma_{o/w} = -\Gamma_s RT d \ln C_s - \Gamma_{co} RT d \ln C_{co} \quad (5.4)$$

Integration of Equation (5.4) gives

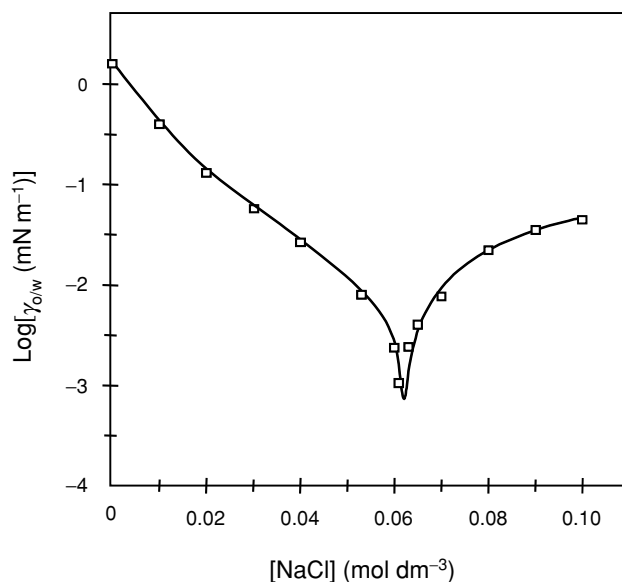
$$\gamma_{o/w} = \gamma_{o/w}^0 - \int_0^{C_s} \Gamma_s RT d \ln C_s - \int_0^{C_{co}} \Gamma_{co} RT d \ln C_{co} \quad (5.5)$$

Equation (5.5) shows that  $\gamma_{o/w}^0$  is lowered by two terms, both from the surfactant and co-surfactant (of surface excesses  $\Gamma_s$  and  $\Gamma_{co}$  respectively) so their effects are additive. It should be mentioned, however, that the two molecules should be adsorbed simultaneously and should not interact with each other (otherwise they lower their respective activities), i.e. are of completely different chemical nature, so that mixed micellisation does not occur.

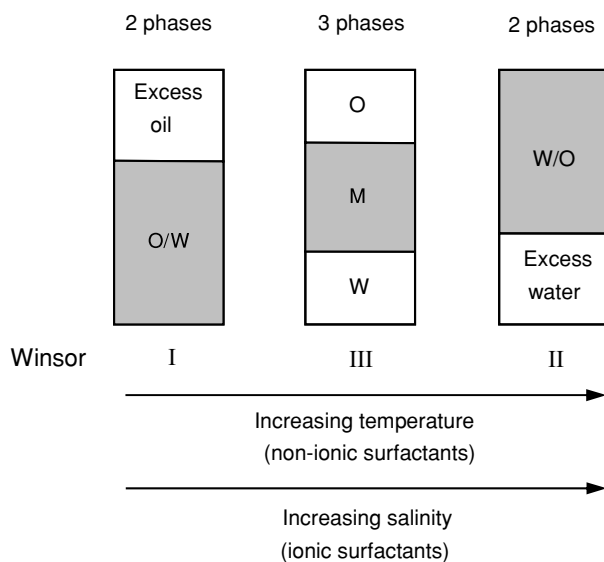
Figure 5.1 shows typical low interfacial tensions found in microemulsions, in this case spanning  $\sim 1 - 10^{-3} \text{ mN m}^{-1}$ . The effect of salt concentration is consistent with changes in the phase behaviour, which are discussed in more detail in Section 5.3 and Figure 5.2.

### 5.2.2 Kinetic instability

Internal contents of the microemulsion droplets are known to exchange, typically on the millisecond time scale [16, 17]. They diffuse and undergo collisions. If collisions are sufficiently violent, then the surfactant film may rupture thereby facilitating droplet exchange, that is the droplets are kinetically unstable. However, if one disperses emulsions as sufficiently small droplets ( $< 500 \text{ \AA}$ ), the tendency to coalesce will be counteracted by an energy barrier. Then the system will remain dispersed and transparent for a long period of time (months) [18]. Such an emulsion is said to be kinetically stable [19]. The mechanism of droplet coalescence has been reported for Aerosol-OT (AOT) w/o microemulsions [16]; the



**Figure 5.1** Oil–water interfacial tension between *n*-heptane and aqueous NaCl solutions as a function of salt concentration in the presence of AOT surfactant. The values were determined by spinning drop tensiometry. The aqueous AOT surfactant concentration is 0.050 mol dm<sup>-3</sup> at temperature 25°C.



**Figure 5.2** Winsor classification and phase sequence of microemulsions encountered as temperature or salinity is scanned for non-ionic and ionic surfactants respectively. Most of the surfactant resides in the shaded area. In the three-phase system the middle-phase microemulsion (M) is in equilibrium with both excess oil (O) and water (W).

droplet exchange process was characterised by a second order rate constant  $k_{\text{ex}}$ , which is believed to be activation controlled (hence the activation energy,  $E_a$ , barrier to fusion) and not purely diffusion controlled. Other studies [20] have shown that the dynamic aspects of microemulsions are affected by the flexibility of the interfacial film, that is film rigidity (see Section 5.3.2), through a significant contribution to the energy barrier. Under the same experimental conditions, different microemulsion systems can have different  $k_{\text{ex}}$  values [16]: for AOT w/o system at room temperature,  $k_{\text{ex}}$  is in the range  $10^6$ – $10^9$   $\text{dm}^3 \text{mol}^{-1} \text{s}^{-1}$ , and for non-ionics  $\text{C}_{12}\text{E}_6$ ,  $10^8$ – $10^9$   $\text{dm}^3 \text{mol}^{-1} \text{s}^{-1}$  [16, 17, 20]. In any case, an equilibrium droplet shape and size are always maintained and this can be studied by different techniques [20].

## 5.3 Physicochemical Properties

This section gives an overview of the main parameters characterising microemulsions. References will be made to related behaviour for planar interfaces presented in Chapter 4.

### 5.3.1 Predicting microemulsion type

A well-known classification of microemulsions is that of Winsor [21] who identified four general types of phase equilibria:

- Type I: the surfactant is preferentially soluble in water and oil-in-water (o/w) microemulsions form (Winsor I). The surfactant-rich water phase coexists with the oil phase where surfactant is only present as monomers at small concentration.
- Type II: the surfactant is mainly in the oil phase and water-in-oil (w/o) microemulsions form. The surfactant-rich oil phase coexists with the surfactant-poor aqueous phase (Winsor II).
- Type III: a three-phase system where a surfactant-rich middle-phase coexists with both excess water and oil surfactant-poor phases (Winsor III or middle-phase microemulsion).
- Type IV: a single-phase (isotropic) micellar solution that forms upon addition of a sufficient quantity of amphiphile (surfactant plus alcohol).

Depending on surfactant type and sample environment, type I, II, III or IV forms preferentially, the dominant type being related to the molecular arrangement at the interface (see below). As illustrated in Figure 5.2, phase transitions are brought about by increasing either electrolyte concentration (in the case of ionic surfactants) or temperature (for non-ionics). Table 5.1 summarises the qualitative changes in phase behaviour of anionic surfactants when formulation variables are modified [22].

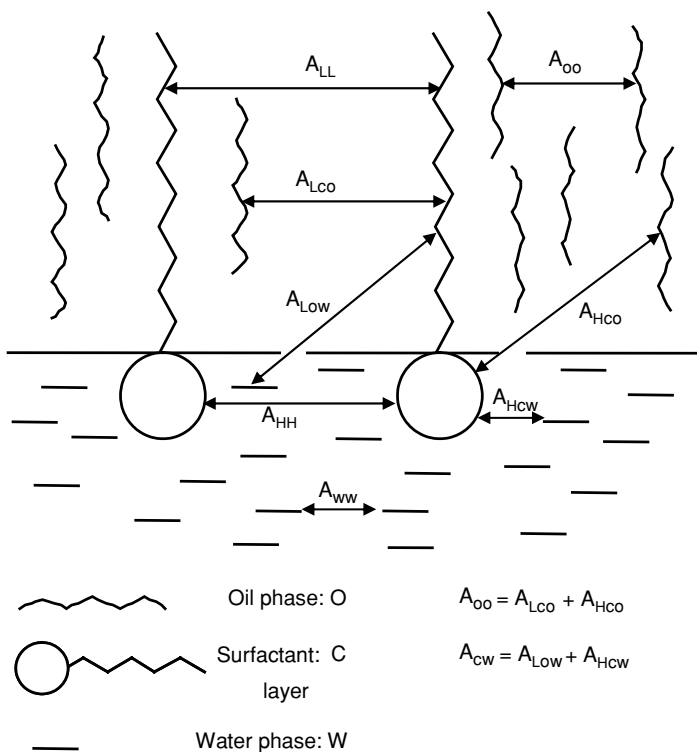
Various investigators have focused on interactions in an adsorbed interfacial film to explain the direction and extent of interfacial curvature. The first concept was that of Bancroft [23] and Clowes [24] who considered the adsorbed film in emulsion systems to be duplex in nature, with an inner and an outer interfacial tension acting independently [25]. The interface would then curve such that the inner surface was one of higher tension. Bancroft's rule was stated as '*that phase will be external in which the emulsifier is most soluble*'; i.e. oil-soluble emulsifiers will form w/o emulsions and water-soluble emulsifiers o/w emulsions. This qualitative concept was largely extended and several parameters have been proposed to quantify the nature of the surfactant film. They are briefly presented in this section. Further

details concerning these microemulsion types and their location in the phase diagram will be given in Section 5.3.3.

*The Winsor R-ratio.* The R-ratio was first proposed by Winsor [21] to account for the influence of amphiphiles and solvents on interfacial curvature. The primary concept is to relate the energies of interaction between the amphiphile layer and the oil and water regions. Therefore, this R-ratio compares the tendency for an amphiphile to disperse into oil to its tendency to dissolve in water. If one phase is favoured, the interfacial region tends to take on a definite curvature. A brief description of the concept is given below, and a full account can be found elsewhere [26].

In micellar or microemulsion solutions, three distinct (single or multi-component) regions can be recognised: an aqueous region, W, an oil or organic region, O, and an amphiphilic region, C. As shown in Figure 5.3, it is useful to consider the interfacial zone as having a definite composition, separating essentially bulk-phase water from bulk-phase oil. In this simple picture, the interfacial zone has a finite thickness, and will contain, in addition to surfactant molecules, some oil and water.

Cohesive interaction energies therefore exist within the C layer, and these determine interfacial film stability. They are depicted schematically in Figure 5.3: the cohesive energy between molecules x and y is defined as  $A_{xy}$ , and is positive whenever interaction between molecules is attractive.  $A_{xy}$  is depicted as the cohesive energy per unit area between surfactant,



**Figure 5.3** Interaction energies in the interfacial region of an oil-surfactant-water system.



**Table 5.1** Qualitative effect of several variables on the observed phase behaviour of anionic surfactants. After Bellocq *et al.* [22]. Concentration effect for low molecular weight alcohols: <sup>a</sup>methanol, ethanol, propanol; <sup>b</sup>higher alkanols

Scanned variables (increase)	Ternary diagram transition
Salinity	I → III → II
Oil: Alkane carbon number	II → III → I
Alcohol: low M.W. <sup>a</sup>	I → III → II
high M.W. <sup>b</sup>	I → III → II
Surfactant: lipophilic chain length	I → III → II
Temperature	II → III → I

oil and water molecules residing in the anisotropic interfacial C layer. For surfactant–oil and surfactant–water interactions  $A_{xy}$  can be considered to be composed of two additive contributions:

$$A_{xy} = A_{Lxy} + A_{Hxy} \quad (5.6)$$

where  $A_{Lxy}$  quantifies interaction between non-polar portions of the two molecules (typically London dispersion forces) and  $A_{Hxy}$  represents polar interactions, especially hydrogen bonding or Coulombic interactions. Thus, for surfactant–oil and surfactant–water interactions, cohesive energies to be considered are

$$A_{co} = A_{Lco} + A_{Hco} \quad (5.7)$$

$$A_{cw} = A_{Lcw} + A_{Hcw} \quad (5.8)$$

$A_{Hco}$  and  $A_{Lcw}$  are generally very small values and can be ignored.

Other cohesive energies are those arising from the following interactions:

- water–water,  $A_{ww}$
- oil–oil,  $A_{oo}$
- hydrophobic–hydrophobic parts (L) of surfactant molecules,  $A_{LL}$
- hydrophilic–hydrophilic parts (H) of surfactant molecules,  $A_{HH}$

The cohesive energy  $A_{co}$  evidently promotes miscibility of the surfactant molecules with the oil region, and  $A_{cw}$  with water. On the other hand,  $A_{oo}$  and  $A_{LL}$  oppose miscibility with oil, while  $A_{ww}$  and  $A_{HH}$  oppose miscibility with water. Therefore, interfacial stability is ensured if the difference in solvent interactions in C with oil and water bulk phases is sufficiently small. Too large a difference, i.e. too strong affinity of C for one phase or the other, would drive to a phase separation.

Winsor expressed qualitatively this variation in dispersing tendency by

$$R = \frac{A_{co}}{A_{cw}} \quad (5.9)$$

To account for the structure of the oil, and the interactions between surfactant molecules, an extended version of the original R-ratio was proposed [26]:

$$R = \frac{(A_{co} - A_{oo} - A_{LL})}{(A_{cw} - A_{ww} - A_{HH})} \quad (5.10)$$

As mentioned before, in many cases,  $A_{\text{Hco}}$  and  $A_{\text{Lcw}}$  are negligible, so  $A_{\text{co}}$  and  $A_{\text{cw}}$  can be approximated respectively to  $A_{\text{Lco}}$  and  $A_{\text{Hcw}}$ .

In brief, Winsor's primary concept is that this R-ratio of cohesive energies, stemming from interaction of the interfacial layer with oil, divided by energies resulting from interactions with water determines the preferred interfacial curvature. Thus, if  $R > 1$ , the interface tends to increase its area of contact with oil while decreasing its area of contact with water. Thus oil tends to become the continuous phase and the corresponding characteristic system is type II (Winsor II). Similarly, a balanced interfacial layer is represented by  $R = 1$ .

*Packing parameter and microemulsion structures.* Changes in film curvature and microemulsion type can be addressed quantitatively in terms of geometric requirements. This concept was introduced by Israelachvili *et al.* [27] and is widely used to relate surfactant molecular structure to interfacial topology. As described in Section 5.3.3, the preferred curvature is governed by relative areas of the head group,  $a_o$ , and the tail group,  $v/l_c$  (see Figure 5.6 for the possible aggregate structures). In terms of microemulsion type,

- if  $a_o > v/l_c$ , then an oil-in-water microemulsion forms
- if  $a_o < v/l_c$ , then a water-in-oil microemulsion forms
- if  $a_o \approx v/l_c$ , then a middle-phase microemulsion is the preferred structure.

*Hydrophilic–lipophilic Balance (HLB).* Another concept relating molecular structure to interfacial packing and film curvature is HLB, the hydrophilic–lipophilic balance. It is generally expressed as an empirical equation based on the relative proportions of hydrophobic and hydrophilic groups within the molecule. The concept was first introduced by Griffin [28] who characterised a number of surfactants, and derived an empirical equation for non-ionic alkyl polyglycol ethers ( $\text{C}_i\text{E}_j$ ) based on the surfactant chemical composition [29]:

$$\text{HLB} = (E_j \text{ wt \%} + \text{OH wt \%})/5 \quad (5.11)$$

where  $E_j$  wt% and OH wt% are the weight per cent of ethylene oxide and hydroxide groups respectively.

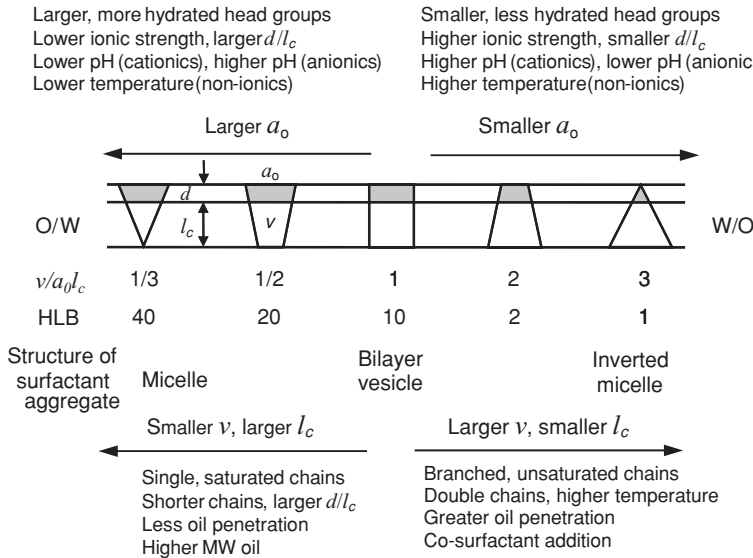
Davies [30] proposed a more general empirical equation that associates a constant to the different hydrophilic and hydrophobic groups:

$$\text{HLB} = [(n_H \times H) - (n_L \times L)] + 7 \quad (5.12)$$

where  $H$  and  $L$  are constants assigned to hydrophilic and hydrophobic groups respectively, and  $n_H$  and  $n_L$  the number of these groups per surfactant molecule.

For bicontinuous structures, i.e. zero curvature, it was shown that  $\text{HLB} \approx 10$  [31]. Then w/o microemulsions form when  $\text{HLB} < 10$ , and o/w microemulsion forms when  $\text{HLB} > 10$ . HLB and packing parameter describe the same basic concept, though the latter is more suitable for microemulsions. The influence of surfactant geometry and system conditions on HLB numbers and packing parameter is illustrated in Figure 5.4.

*Phase Inversion Temperature (PIT).* Non-ionic surfactants form water–oil microemulsions (and emulsions) with a high temperature sensitivity. In particular, there is a specific



**Figure 5.4** Effect of molecular geometry and system conditions on the packing parameter and HLB number (reprinted from Israelachvili [31] with permission from Elsevier).

PIT and the film curvature changes from positive to negative. This critical point was defined by Shinoda and Saito [32]:

- if  $T < \text{PIT}$ , an oil-in-water microemulsion forms (Winsor I)
- if  $T > \text{PIT}$ , a water-in-oil microemulsion forms (Winsor II)
- at  $T = \text{PIT}$ , a middle-phase microemulsion exists (Winsor III) with a spontaneous curvature equal to zero, and a HLB number (Equation (5.11)) approximately equal to 10.

The HLB number and PIT are therefore connected; hence the term HLB temperature is sometimes employed [33].

### 5.3.2 Surfactant film properties

An alternative, more physically intuitive, approach is to consider mechanical properties of a surfactant film at an oil–water interface. This film can be characterised by three phenomenological constants: tension, bending rigidity and spontaneous curvature. Their relative importance depends on the constraints felt by the film. It is important to understand how these parameters relate to interfacial stability since surfactant films determine the static and dynamic properties of microemulsions (and emulsions). These include phase behaviour and stability, structure and solubilisation capacity.

*Ultra-low interfacial tension.* Interfacial (or surface) tensions,  $\gamma$ , were defined in Section 4.3.1 for planar surfaces, and the same principle applies for curved liquid–liquid interfaces, i.e. it corresponds to the work required to increase interfacial area by unit amount. As mentioned in Section 5.2.1, microemulsion formation is accompanied by ultra-low interfacial

oil–water tensions,  $\gamma_{o/w}$ , typically  $10^{-2} - 10^{-4} \text{ mN m}^{-1}$ . They are affected by the presence of a co-surfactant, as well as electrolyte and/or temperature, pressure and oil chain length. Several studies have been reported on the effect of such variables on  $\gamma_{o/w}$ . In particular, Aveyard and co-workers performed several systematic interfacial tension measurements on both ionic [34, 35] and non-ionics [36], varying oil chain length, temperature and electrolyte content. For example, as depicted in Figure 5.1, in the system water–AOT–*n*-heptane, at constant surfactant concentration (above its CMC), a plot of  $\gamma_{o/w}$  as a function of electrolyte (NaCl) concentration shows a deep minimum that corresponds to the Winsor phase inversion; i.e. upon addition of NaCl,  $\gamma_{o/w}$  decreases to a minimum critical value (Winsor III structure), then increases to a limiting value close to  $0.2\text{--}0.3 \text{ mN m}^{-1}$  (Winsor II region). At constant electrolyte concentration, varying temperature [34], oil chain length and co-surfactant content [35] have a similar effect. With non-ionics, a similar tension curve and phase inversion are observed, but on increasing temperature [36]. In addition, when increasing surfactant chain length, the interfacial tension curves shift to higher temperatures and the minimum in  $\gamma_{o/w}$  decreases [37]. Ultra-low interfacial tensions cannot be measured with standard techniques such as Du Nouy Ring, Wilhelmy plate or drop volume (DVT). Appropriate techniques for this low tension range are spinning drop tensiometry (SDT) and surface light scattering [38].

*Spontaneous curvature.* Spontaneous (or natural or preferred) curvature  $C_o$  is defined as the curvature formed by a surfactant film when a system consists of equal amounts of water and oil. Then, there is no constraint on the film, which is free to adopt the lowest free energy state. Whenever one phase is predominant, there is a deviation from  $C_o$ . In principle, every point on a surface possesses two principal radii of curvature  $R_1$  and  $R_2$  and their associated principal curvatures are  $C_1 = 1/R_1$  and  $C_2 = 1/R_2$ . Mean and Gaussian curvatures are used to define the bending of surfaces. They are defined as follows [39]:

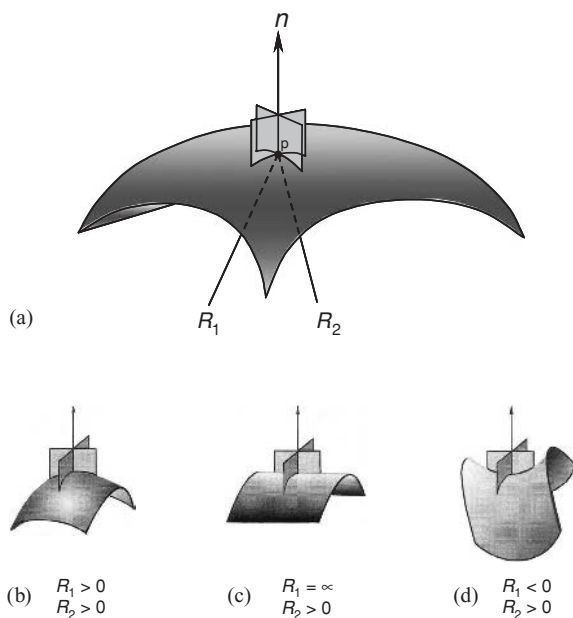
$$\text{Mean curvature : } C = (1/2)(1/R_1 + 1/R_2)$$

$$\text{Gaussian curvature : } \kappa = 1/R_1 \times 1/R_2$$

$C_1$  and  $C_2$  are determined as follows: every point on the surface of the surfactant film has two principal radii of curvature  $R_1$  and  $R_2$  as shown in Figure 5.5. If a circle is placed tangentially to a point  $p$  on the surface and if the circle radius is chosen so that its second derivative at the contact point equals that of the surface in the direction of the tangent (of normal vector,  $n$ ), then the radius of the circle is a radius of curvature of the surface. The curvature of the surface is described by two such circles chosen in orthogonal (principal) directions as shown in Figure 5.5a.

For a sphere,  $R_1$  and  $R_2$  are equal and positive (Figure 5.5b). For a cylinder  $R_2$  is indefinite (Figure 5.5c) and for a plane, both  $R_1$  and  $R_2$  are indefinite. In the special case of a saddle,  $R_1 = -R_2$ , i.e. at every point the surface is both concave and convex (Figure 5.5d). Both a plane and saddle have the property of zero mean curvature.

The curvature  $C_o$  depends both on the composition of the phases it separates and on surfactant type. One argument applied to the apolar side of the interface is that oil can penetrate to some extent between the surfactant hydrocarbon tails. The more extensive the penetration, the more curvature is imposed towards the polar side. This results in a decrease of  $C_o$  since, by convention, positive curvature is towards oil (and negative towards



**Figure 5.5** Principal curvatures of different surfaces. (a) Intersection of the surfactant film surface with planes containing the normal vector ( $n$ ) to the surface at the point  $p$ . (b) convex curvature, (c) cylindrical curvature and (d) saddle-shaped curvature. Reprinted from Hyde *et al.*, Chapter 1, Figure 1.5 [39], with permission from Elsevier.

water). The longer the oil chains, the less they penetrate the surfactant film and the smaller the effect on  $C_o$ . Recently, Eastoe *et al.* have studied the extent of solvent penetration in microemulsions stabilised by di-chained surfactants, using SANS and selective deuteration. Results suggested that oil penetration is a subtle effect, which depends on the chemical structures of both surfactant and oil. In particular, unequal surfactant chain length [40–43] or presence of C=C bonds [44] results in a more disordered surfactant/oil interface, thereby providing a region of enhanced oil mixing. For symmetric di-chained surfactants (e.g. DDAB and AOT), however, no evidence for oil mixing was found [42]. The effects of alkane structure and molecular volume on the oil penetration were also investigated with *n*-heptane, and cyclohexane. The results indicate that heptane is essentially absent from the layers, but the more compact molecule cyclohexane has a greater penetrating effect [43].

Surfactant type, and nature of the polar head group, also influences  $C_o$  through different interactions with the polar (aqueous) phase:

- For ionic surfactants, electrolyte content and temperature affect the spontaneous curvature in opposite ways. An increase in salt concentration screens electrostatic head group repulsions – i.e. decreases head group area – so the film curves more easily towards water, leading to a decrease in  $C_o$ . Raising the temperature has two effects: (1) an increase in electrostatic repulsions between head groups due to higher counterion dissociation, so  $C_o$  increases and (2) more gauche conformations are induced in the surfactant chains, which become more coiled, resulting in a decrease in  $C_o$ . Therefore the combined effects of temperature on the apolar chains and on electrostatic interactions are competitive.

The electrostatic term is believed to be slightly dominant, so  $C_o$  increases weakly with increasing temperature.

- For non-ionic surfactants, unsurprisingly, electrolytes have very little effect on  $C_o$ , whereas temperature is a critical parameter due to the strong dependence of their solubility (in water or oil) on temperature. For surfactants of the  $C_iE_j$  type as temperature increases water becomes a less good solvent for the hydrophilic units and penetrates less into the surfactant layer. In addition, on the other side of the film, oil can penetrate further into the hydrocarbon chains, so that increasing the temperature for this type of surfactant causes a strong decrease in  $C_o$ . This phenomenon explains the strong temperature effects on the phase equilibria of such surfactants as shown in Figure 5.8 (see Section 4.4).

Thus, by changing external parameters such as temperature, nature of the oil or electrolyte concentration, the spontaneous curvature can be tuned to the appropriate value, and so drive transitions between Winsor systems. Other factors affect  $C_o$  in a similar fashion; they include variation of the polar head group, type and valency of counterions, length and number of apolar chains, addition of a co-surfactant or mixing of surfactants.

*Film bending rigidity.* Film rigidity is an important parameter associated with interfacial curvature. The concept of film bending energy was first introduced by Helfrich [45] and is now considered as an essential model for understanding microemulsion properties. It can be described by two elastic moduli [46] that measure the energy required to deform the interfacial film from a preferred mean curvature:

- the mean bending elasticity (or rigidity),  $K$ , associated with the mean curvature, that represents the energy required to bend unit area of surface by unit amount.  $K$  is positive, i.e. spontaneous curvature is favoured
- the more elusive factor  $\bar{K}$  is associated with Gaussian curvature, and hence accounts for film topology.  $\bar{K}$  is negative for spherical structures or positive for bicontinuous cubic phases.

Theoretically, it is expected that bending moduli should depend on surfactant chain length [47], area per surfactant molecule in the film [48] and electrostatic head group interactions [49].

The film rigidity theory is based on the interfacial free energy associated with film curvature. The free energy,  $F$ , of a surfactant layer at a liquid interface may be given by the sum of an interfacial energy term,  $F_i$ , a bending energy term,  $F_b$ , and an entropic term,  $F_{ent}$ . For a droplet type structure this is written as [50]

$$F = F_i + F_b + F_{ent} = \gamma A + \int \left[ \frac{K}{2} (C_1 + C_2 - 2C_o)^2 + \bar{K} C_1 C_2 \right] dA + nk_B T f(\phi) \quad (5.13)$$

where  $\gamma$  is the interfacial tension,  $A$  is the total surface area of the film,  $K$  is the mean elastic bending modulus,  $\bar{K}$  is the Gaussian bending modulus,  $C_1$  and  $C_2$  are the two principal curvatures,  $C_o$  is the spontaneous curvature,  $n$  is the number of droplets,  $k_B$  is the Boltzmann constant and  $f(\phi)$  is a function accounting for the entropy of mixing of the microemulsion droplets, where  $\phi$  is the droplet core volume fraction. For dilute systems where  $\phi < 0.1$ , it was shown that  $f(\phi) = [\ln(\phi) - 1]$  [50]. Microemulsion formation is associated with

ultra-low interfacial tension  $\gamma$  so the  $\gamma A$  term is small compared to  $F_b$  and  $F_{\text{ent}}$ , and can be ignored as an approximation.

As mentioned previously, the curvatures  $C_1$ ,  $C_2$  and  $C_o$  can be expressed in terms of radii as  $1/R_1$ ,  $1/R_2$  and  $1/R_o$  respectively. For spherical droplets,  $R_1 = R_2 = R$ , and the interfacial area is  $A = n4\pi R^2$ . Note that  $R$  and  $R_o$  are core radii rather than droplet radii [50]. Solving Equation (5.13) and dividing by area  $A$ , the total free energy,  $F$ , for spherical droplets (of radius  $R$ ) is expressed as

$$\frac{F}{A} = 2K \left( \frac{1}{R} - \frac{1}{R_o} \right)^2 + \frac{\bar{K}}{R^2} + \left[ \frac{k_B T}{4\pi R^2} f(\phi) \right] \quad (5.14)$$

For systems where the solubilisation boundary is reached (WI or WII region), a microemulsion is in equilibrium with an excess phase of the solubilisate and the droplets have achieved their maximum size, i.e. the maximum core radius,  $R_{\text{max}}^{\text{av}}$ . Under this condition the minimisation of the total free energy leads to a relation between the spontaneous radius,  $R_o$ , and the elastic constants  $K$  and  $\bar{K}$  [51]:

$$\frac{R_{\text{max}}^{\text{av}}}{R_o} = \frac{2K + \bar{K}}{2K} + \frac{k_B T}{8\pi K} f(\phi) \quad (5.15)$$

A number of techniques have been used to determine  $K$  and  $\bar{K}$  separately, in particular, ellipsometry, X-ray reflectivity and small-angle X-ray scattering (SAXS) techniques [52–54]. De Gennes and Taupin [55] have developed a model for bicontinuous microemulsions. For  $C_o = 0$  the layer is supposed to be flat in the absence of thermal fluctuations. They introduced the term  $\xi_K$ , the persistence length of the surfactant layer that relates to  $K$ , via

$$\xi_K = a \exp(2\pi K / k_B T) \quad (5.16)$$

where  $a$  is a molecular length and  $\xi_K$  is the correlation length for the normals to the layer, i.e. the distance over which this layer remains flat in the presence of thermal fluctuations.  $\xi_K$  is extremely sensitive to the magnitude of  $K$ . When  $K \gg k_B T$ ,  $\xi_K$  is macroscopic, i.e. the surfactant layer is flat over large distances and ordered structures such as lamellar phases may form. If  $K$  is reduced to  $\sim k_B T$  then  $\xi_K$  is microscopic, ordered structures are unstable and disordered phases such as microemulsions may form. Experiments reveal that  $K$  is typically between  $100k_B T$  for condensed insoluble monolayers [56] and about  $10k_B T$  for lipid bilayers [57–59] but can decrease below  $k_B T$  in microemulsion systems [60]. The role of  $\bar{K}$  is also important; however, there are few measurements of this quantity in the literature [e.g. 53, 61]. Its importance in determining the structure of surfactant–oil–water mixtures is still far from clear.

An alternative, more accessible, method to quantify film rigidities is to calculate the composite parameter  $(2K + \bar{K})$  using tensiometry and SANS techniques. This parameter can be derived for droplet microemulsion at the solubilisation boundary, WI or WII system, by combining the radius of the droplet with interfacial tensions or droplet polydispersity. Two expressions can be derived from Equations (5.17) and (5.18).

1. *Using the interfacial tension  $\gamma_{o/w}$  (measured by SLS or SDT) and the maximum mean core radius  $R_{\text{max}}^{\text{av}}$  (measured by SANS, see Chapter 12):*  $\gamma_{o/w}$  at the interface between microemulsion and excess phases at the solubilisation boundary can be expressed in terms of these elastic moduli and  $R_{\text{max}}^{\text{av}}$  [52]. Any new area created must be covered by a

monolayer of surfactant, and so this energy may be calculated in the case of WI or WII systems since the surfactant monolayer is taken from around the curved microemulsion droplets [56]. To do this it is necessary to unbend the surfactant film, introducing a contribution from  $K$ , of magnitude  $2K/(R_{\max}^{\text{av}})^2$ . The resulting change in the number of microemulsion droplets introduces an entropic contribution and a contribution due to the change in topology involving  $\bar{K}$ , of magnitude  $\bar{K}/(R_{\max}^{\text{av}})^2$ . So the interfacial tension between the microemulsion and excess phase is given by

$$\gamma_{\text{o/w}} = \frac{2K + \bar{K}}{(R_{\max}^{\text{av}})^2} + \frac{k_{\text{B}} T}{4\pi (R_{\max}^{\text{av}})^2} f(\phi) \quad (5.17)$$

which gives for the bending moduli:

$$2K + \bar{K} = \gamma_{\text{o/w}} (R_{\max}^{\text{av}})^2 - \frac{k_{\text{B}} T}{4\pi} f(\phi) \quad (5.18)$$

2. *Using the width of the Schultz polydispersity function  $p = \sigma/R_{\max}^{\text{av}}$  obtained from SANS analysis (see Section 12.16):* droplet polydispersity relates to the bending moduli through thermal fluctuations of the microemulsion droplets. Safran [62] and Milner [63] described the thermal fluctuations by an expansion of the droplet deformation in terms of spherical harmonics [50]. In the case of the two-phase equilibria at maximum solubilisation (WI or WII), this polydispersity,  $p$ , may be expressed as a function of  $K$  and  $\bar{K}$ :

$$p^2 = \frac{u_0^2}{4\pi} = \frac{k_{\text{B}} T}{8\pi (2K + \bar{K}) + 2k_{\text{B}} T f(\phi)} \quad (5.19)$$

where  $u_0$  is the fluctuation amplitude for the  $l = 0$  spherical dispersity mode. This polydispersity is given by the SANS Schultz polydispersity parameter  $\sigma/R_{\max}^{\text{av}}$  [64], and Equation (5.19) can be written as

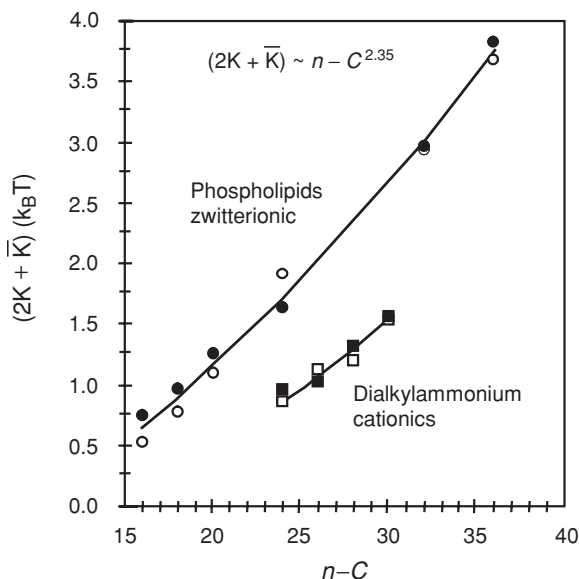
$$2K + \bar{K} = \frac{k_{\text{B}} T}{8\pi (\sigma/R_{\max}^{\text{av}})^2} - \frac{k_{\text{B}} T}{4\pi} f(\phi) \quad (5.20)$$

Therefore Equations (5.17) and (5.20) give two accessible expressions for the sum  $(2K + \bar{K})$  using data from SANS and tensiometry. This approach has been shown to work well with non-ionic films in WI systems [50, 65], and also cationic [64] and zwitterionic [66] layers in WII microemulsions. Figure 5.6 shows results for these latter two classes of system, as a function of surfactant alkyl carbon number  $n$ -C. The good agreement between Equations (5.17) and (5.20) suggest they can be used with confidence. These values are in line with current statistical mechanical theories [48], which suggest that  $K$  should vary as  $n\text{-C}^{2.5}$  to  $n\text{-C}^3$ , whereas there is only a small effect on  $\bar{K}$ .

### 5.3.3 Microemulsion phase behaviour

Solubilisation and interfacial properties of microemulsions depend upon pressure, temperature and also on the nature and concentration of the components. The determination of phase stability diagrams (or phase maps), and location of the different structures formed within these water(salt)–oil–surfactant–alcohol systems in terms of variables are, therefore,





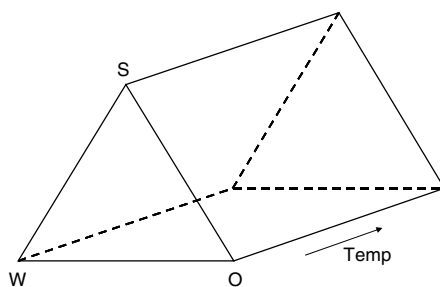
**Figure 5.6** Film rigidities ( $2K + \bar{K}$ ) as a function of total alkyl carbon number  $n-C$  from Winsor II microemulsions. The lines are guides to the eye.

very important. Several types of phase diagrams can be identified depending on the number of variables involved. In using an adequate mode of representation, it is possible not only to describe the limits of existence of the single and multi-phase regions, but also to characterise equilibria between phases (tie-lines, tie-triangles, critical points, etc.). Given below is a brief description of ternary and binary phase maps, as well as the phase rule that dictates their construction.

*Phase rule.* The phase rule enables the identification of the number of variables (or degrees of freedom) depending on the system composition and conditions. It is generally written as [67]:

$$F = C - P + 2 \quad (5.21)$$

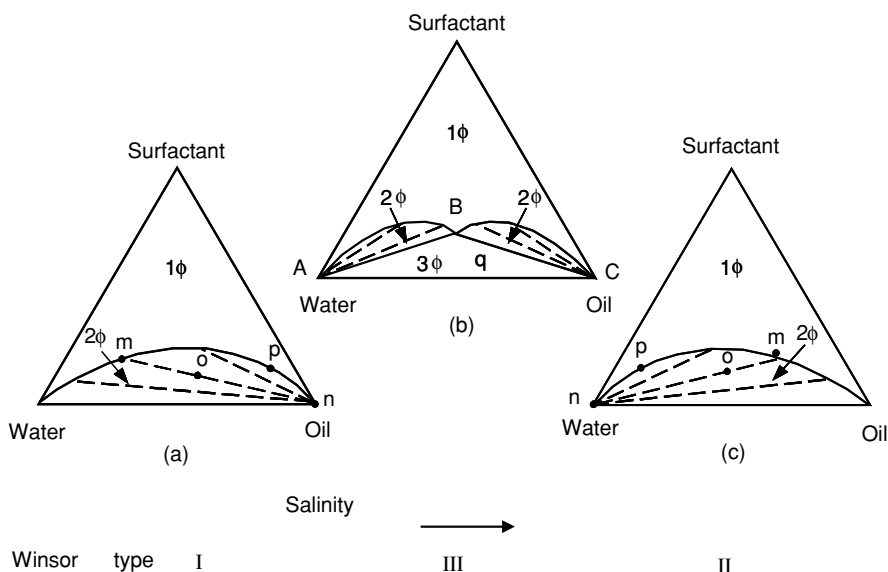
where  $F$  is the number of possible independent changes of state or degrees of freedom,  $C$  the number of independent chemical constituents and  $P$  the number of phases present in the system. A system is called invariant, monovariant, bivariant and so on, according to whether  $F$  is 0, 1, 2 and so on. For example, in the simplest case of a system composed of three components and two phases,  $F$  is univariant at a fixed temperature and pressure. This means that the mole or weight fraction of one component in one of the phases can be specified but all other compositions in both phases are fixed. In general, microemulsions contain at least three components: oil (O), water (W) and amphiphile (S), and as mentioned previously a co-surfactant (alcohol) and/or an electrolyte are usually added to tune the system stability. These can be considered as simple O–W–S systems: whenever a co-surfactant is used, the ratio oil:alcohol is usually kept constant and it is assumed that the alcohol does not interact



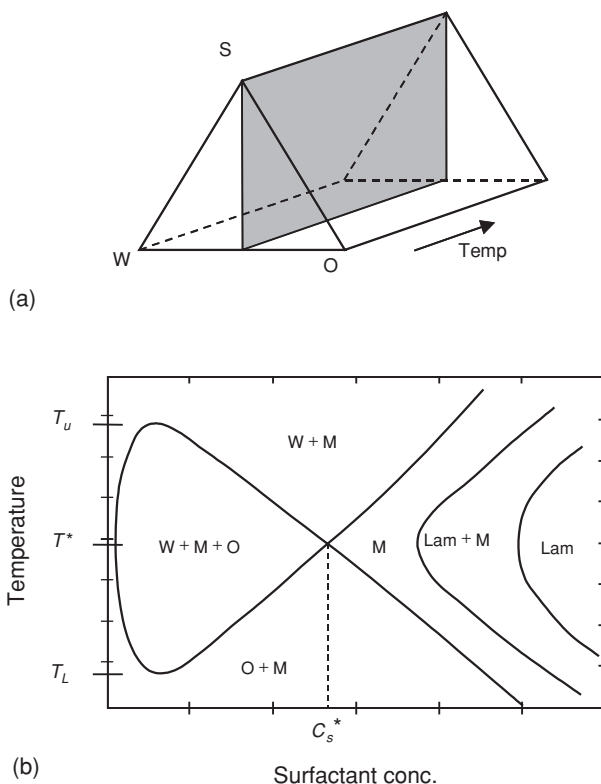
**Figure 5.7** The phase prism, describing the phase behaviour of a ternary system at constant pressure.

with any other component so that the mixture can be treated (to a first approximation) as a three-component system. At constant pressure, the composition–temperature phase behaviour can be presented in terms of a phase prism, as illustrated in Figure 5.7. However, the construction of such a phase map is rather complex and time consuming so it is often convenient to simplify the system by studying specific phase-cuts. The number of variables can be reduced either by keeping one term constant and/or by combining two or more variables. Then, ternary and binary phase diagrams are produced.

*Ternary phase diagrams.* At constant temperature and pressure, the ternary phase diagram of a simple three-component microemulsion is divided into two or four regions as shown in Figure 5.8. In each case, every composition point within the single-phase region above



**Figure 5.8** Ternary diagram representations of two- and three-phase regions formed by simple water–oil–surfactant systems at constant temperature and pressure. (a) Winsor I type, (b) Winsor II type, (c) Winsor III type systems.

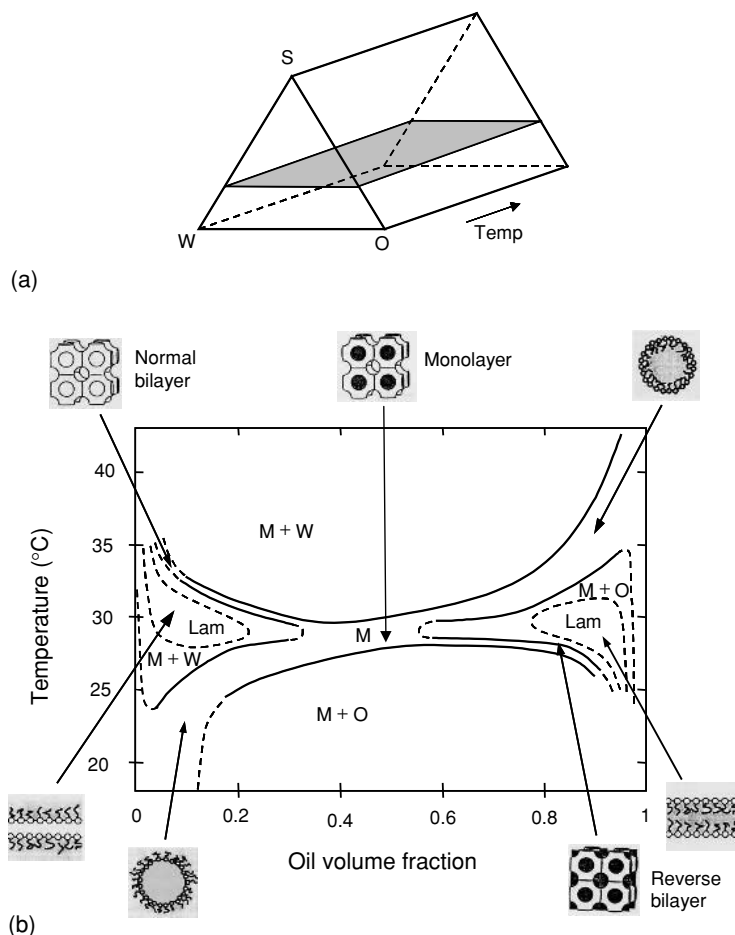


**Figure 5.9** Binary phase behaviour in ternary microemulsion systems formed with non-ionic surfactants. (a) Illustration of the section through the phase prism at equal water and oil contents. (b) Schematic phase diagram plotted as temperature versus surfactant concentration  $C_s$ .  $T_L$  and  $T_U$  are the lower and upper temperatures, respectively, of the phase equilibrium  $W + M + O$ .  $T^*$  is the temperature at which the three-phase triangle is an isosceles, i.e. when the middle-phase microemulsion contains equal amounts of water and oil. This condition is also termed 'balanced'.  $C_s^*$  is the surfactant concentration in the middle-phase microemulsion at balanced conditions. 'Lam' denotes a lamellar liquid crystalline phase. Reprinted from Olsson and Wennerström [68] with permission from Elsevier.

the demixing line corresponds to a microemulsion. Composition points below this line correspond to multi-phase regions comprising in general microemulsions in equilibrium with either an aqueous or an organic phase or both, i.e. Winsor-type systems (see Section 5.3.1).

Any system whose overall composition lies within the two-phase region (e.g. point o in Figures 5.8a and 5.8c) will exist as two phases whose compositions are represented by the ends of the 'tie-line', i.e. a segment formed by phases m and n. Therefore, every point on a particular tie-line has identical coexisting phases (m and n) but of different relative volumes. When the two conjugate phases have the same composition ( $m = n$ ), this corresponds to the plait (or critical) point, p.

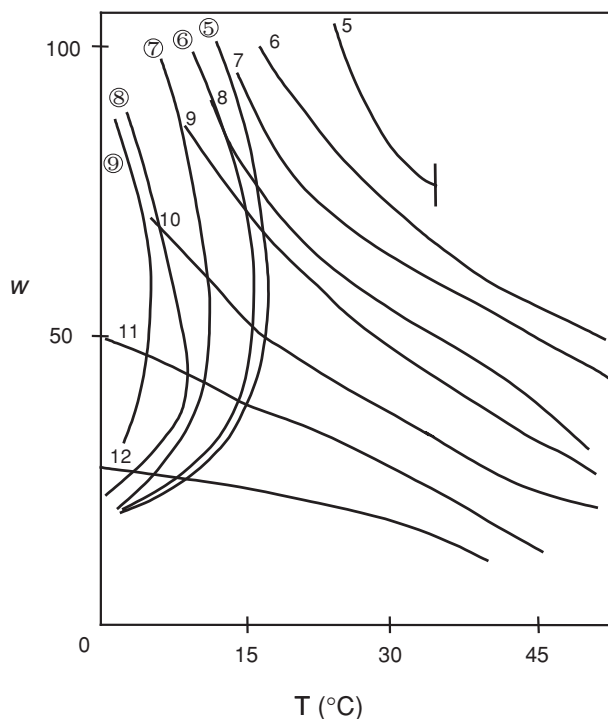
If three phases coexist (Figure 5.8b), i.e. corresponding to WIII, the system at constant temperature and pressure is, according to the phase rule, invariant. Then, there is a region of the ternary diagram that consists of three-phase systems having invariant compositions and whose boundaries are tie-lines in the adjacent two-phase regions that surround it. This



**Figure 5.10** Binary phase behaviour in ternary microemulsion systems formed with non-ionic surfactants. (a) Illustration of a section at constant surfactant concentration through the phase prism. (b) Schematic phase diagram, plotted as temperature versus volume fraction of oil,  $\phi_o$ , at constant surfactant concentration. Also shown are various microstructures found in different regions of the microemulsion phase, M. At higher temperatures the liquid phase is in equilibrium with excess water (M + W), and at lower temperatures with excess oil (M + O). At intermediate temperatures a lamellar phase is stable at higher water contents and higher oil contents, respectively. Reprinted from Olsson and Wennerström [68] with permission from Elsevier.

region of three-phase invariant compositions is therefore triangular in form and called 'tie-triangle' [26]. Any overall composition, such as point q (Figure 5.8b) lying within the tie-triangle, will divide into three phases having compositions corresponding to the vertices A, B and C of the triangle. The compositions A, B and C are invariant in the sense that varying the position q, the overall composition throughout the triangle will result in variations in the amounts of the phases A, B and C but not in their composition.

*Binary phase diagrams.* As mentioned previously, ternary diagrams can be further simplified by fixing some parameters and/or combining two variables together (e.g. water and



**Figure 5.11** Pseudo-binary phase diagram in ternary microemulsion systems formed with the anionic surfactant AOT in various straight-chain alkane solvents. The water-to-surfactant molar ratio,  $w$ , is plotted versus temperature at constant surfactant concentration and pressure. Alkane carbon numbers are indicated; circled numbers correspond to the lower temperature (solubilisation) boundary,  $T_{low}$ , and uncircled numbers to the upper temperature (haze) boundary,  $T_{high}$ . The single phase microemulsion region is located between  $T_{low}$  and  $T_{high}$ . Below  $T_{low}$  the system consists of a microemulsion phase in equilibrium with excess water (WII type), and above  $T_{high}$  the single microemulsion phase separates into a surfactant-rich phase and an oil phase. Fletcher *et al.* [16], reproduced by permission of The Royal Society of Chemistry.

electrolyte into brine, or water and oil into water-to-oil ratio), i.e. reducing the degrees of freedom. Then, determining the phase diagram of such systems reduces to a study of a planar section through the phase prism. Examples of such pseudo-binary diagrams are given in Figures 5.9–5.11 for non-ionic and anionic surfactants.

Figure 5.9 shows the schematic phase diagram for a non-ionic surfactant–water–oil ternary system. Since temperature is a crucial variable in the case of non-ionics, the pseudo-binary diagram is represented by the planar section defined by  $\phi_w = \phi_o$ , where  $\phi_w$  and  $\phi_o$  are the volume fractions of water and oil respectively. Then, at constant pressure, defining the system in a single-phase region requires the identification of two independent variables ( $F = 2$ ), i.e. temperature and surfactant concentration. The points  $T_L$  and  $T_U$  represent the lower and upper temperatures, respectively, of the phase equilibrium  $W + M + O$  (with  $M$  being the microemulsion phase). The minimum amount of surfactant necessary to solubilise equal amounts of water and oil is denoted  $C_s^*$  [68]. The lower  $C_s^*$  the more efficient the surfactant. Figure 5.10 illustrates the determination of a second possible section for a non-ionic surfactant–water–oil ternary system: pressure and surfactant concentration are kept constant, leaving the two variables, temperature and water-to-oil ratio ( $\phi_{w-o}$ ).

This diagram shows the various surfactant phases obtained as a function of temperature and water-to-oil ratio [68]. The third example (Figure 5.11) concerns an anionic surfactant, Aerosol-OT. In order to obtain  $F = 2$  when defining the ternary W–O–S system in a single-phase region at constant pressure, the surfactant concentration parameter is fixed. Then, the two variables are temperature and  $w$ , the water-to-surfactant molar ratio defined as  $w = [\text{water}]/[\text{surfactant}]$ .  $w$  represents the number of water molecules solubilised per surfactant molecule, so that this phase diagram characterises the surfactant efficiency, as a microemulsifier.

## References

1. Danielsson, I. and Lindman, B. (1981) *Colloids Surf. A*, **3**, 391.
2. Sjöblom, J., Lindberg, R. and Friberg, S.E. (1996) *Adv. Colloid Interface Sci.*, **65**, 125
3. Schulman, J.H., Stoeckenius, W. and Prince, M. (1959) *J. Phys. Chem.*, **63**, 1677.
4. Shinoda, K. and Friberg, S. (1975) *Adv. Colloid Interface Sci.*, **4**, 281.
5. Adamson, A.W. (1969) *J. Colloid Interface Sci.*, **29**, 261.3.
6. Friberg, S.E., Mandell, L. and Larsson, M. (1969) *J. Colloid Interface Sci.*, **29**, 155.
7. Shah, D.O. (ed.) (1981) *Surface Phenomena in Enhanced Recovery*. Plenum, New York.
8. Overbeek, J.Th.G. (1978) *Faraday Discuss. Chem. Soc.*, **65**, 7.
9. Tadros, Th.F. and Vincent, B. (1980) In P. Becher (ed.), *Encyclopaedia of Emulsion Technology*. Dekker, New York. vol. 1.
10. Kunieda, H. and Shinoda, K. (1980) *J. Colloid Interface Sci.*, **75**, 601.
11. Chen, S.J., Evans, F.D. and Ninham, B.W. (1984) *J. Phys. Chem.*, **88**, 1631.
12. Kahlweit, M., Strey, R. and Busse, G. (1990) *J. Phys. Chem.*, **94**, 3881.
13. Hunter, R.J. (1994) *Introduction to Modern Colloid Science*. Oxford University Press, Oxford.
14. Lekkerkerker, H.N.W., Kegel, W.K. and Overbeek, J.Th.G. (1996) *Ber. Bunsenges Phys. Chem.*, **100**, 206.
15. Ruckenstein, E. and Chi, J.C. (1975) *J. Chem. Soc. Faraday Trans.*, **71**, 1690.
16. Fletcher, P.D.I., Howe, A.M. and Robinson, B.H. (1987) *J. Chem. Soc. Faraday Trans. 1*, **83**, 985.
17. Fletcher, P.D.I., Clarke, S. and Ye, X. (1990) *Langmuir*, **6**, 1301.
18. Biais, J., Bothorel, P., Clin, B. and Lalanne, P. (1981) *J. Colloid Interface Sci.*, **80**, 136.
19. Friberg, S., Mandell, L. and Larson, M. (1969) *J. Colloid Interface Sci.*, **29**, 155.
20. Fletcher, P.D.I. and Horsup, D.I. (1992) *J. Chem. Soc. Faraday Trans. 1*, **88**, 855.
21. Winsor, P.A. (1948) *Trans. Faraday Soc.*, **44**, 376.
22. Bellocq, A.M., Biais, J., Bothorel, P., Clin, B., Fourche, G., Lalanne, P., Lemaire, B., Lemanceau, B. and Roux, D. (1984) *Adv. Colloid Interface Sci.*, **20**, 167.
23. Bancroft, W.D. (1913) *J. Phys. Chem.*, **17**, 501.
24. Clowes, G.H.A. (1916) *J. Phys. Chem.*, **20**, 407.
25. Adamson, A.W. (1960) *Physical Chemistry of Surfaces*. Interscience, New York. p. 393.
26. Bourrel, M. and Schechter, R.S. (1988) *Microemulsions and Related Systems*. Dekker, New York.
27. Israelachvili, J.N., Mitchell, D.J. and Ninham, B.W. (1976) *J. Chem. Soc. Faraday Trans. 2*, **72**, 1525.
28. Griffin, W.C. (1949) *J. Cosmetics Chem.*, **1**, 311.
29. Griffin, W.C. (1954) *J. Cosmetics Chem.*, **5**, 249.
30. Davies, J.T. (1959) *Proc. 2nd Int. Congr. Surface Act.* Butterworths, London. vol. 1.
31. Israelachvili, J.N. (1994) *Colloids Surf. A*, **91**, 1.
32. Shinoda, K. and Saito, H. (1969) *J. Colloid Interface Sci.*, **34**, 238.
33. Shinoda, K. and Kunieda, H. (1983) In P. Becher (ed.) *Encyclopaedia of Emulsion Technology*. Dekker, New York. vol. 1.

34. Aveyard, R., Binks, B.P., Clarke, S. and Mead, J. (1986) *J. Chem. Soc. Faraday Trans. 1*, **82**, 125.
35. Aveyard, R., Binks, B.P. and Mead, J. (1986) *J. Chem. Soc. Faraday Trans. 1*, **82**, 1755.
36. Aveyard, R., Binks, B.P. and Fletcher, P.D.I. (1989) *Langmuir*, **5**, 1210.
37. Sottmann, T. and Strey, R. (1996) *Ber. Bunsenges Phys. Chem.*, **100**, 237.
38. Langevin, D. (ed.) (1992) *Light Scattering by Liquid Surfaces and Complementary Techniques*. Dekker, New York.
39. Hyde, S., Andersson, K., Larsson, K., Blum, Z., Landh, S. and Ninham, B.W. (1997) *The Language of Shape*. Elsevier, Amsterdam.
40. Eastoe, J., Dong, J., Hetherington, K.J., Steytler, D.C. and Heenan, R.K. (1996) *J. Chem. Soc. Faraday Trans.*, **92**, 65.
41. Eastoe, J., Hetherington, K.J., Sharpe, D., Dong, J., Heenan, R.K. and Steytler, D.C. (1996) *Langmuir*, **12**, 3876.
42. Eastoe, J., Hetherington, K.J., Sharpe, D., Dong, J., Heenan, R.K. and Steytler, D.C. (1997) *Colloids Surf. A*, **128**, 209.
43. Eastoe, J., Hetherington, K.J., Sharpe, D., Steytler, D.C., Egelhaaf, S. and Heenan, R.K. (1997) *Langmuir*, **13**, 2490.
44. Bumajdad, A., Eastoe, J., Heenan, R.K., Lu, J.R., Steytler, D.C. and Egelhaaf, S. (1998) *J. Chem. Soc. Faraday Trans.*, **94**, 2143.
45. Helfrich, W.Z. (1973) *Naturforsch.*, **28c**, 693.
46. Kellay, H., Binks, B.P., Hendriks, Y., Lee, L.T. and Meunier, J. (1994) *Adv. Colloid Interface Sci.*, **9**, 85.
47. Safran, S.A. and Tlusty, T. (1996) *Ber. Bunsenges. Phys. Chem.*, **100**, 252.
48. Szleifer, I., Kramer, D., Ben-Shaul, A., Gelbart, W.M. and Safran, S. (1990) *J. Chem. Phys.*, **92**, 6800.
49. Winterhalter, M. and Helfrich, W. (1992) *J. Phys. Chem.*, **96**, 327.
50. Gradzielski, M., Langevin, D. and Farago, B. (1996) *Phys. Rev. E*, **53**, 3900.
51. Safran, S.A. (1992) In S.H. Chen, J.S. Huang and P. Tartaglia (eds) *Structure and Dynamics of Strongly Interacting Colloids and Supramolecular Aggregates in Solution*. Vol. 369 of NATO Advanced Study Institute, Series C: Mathematical and Physical Sciences. Kluwer, Dordrecht.
52. Meunier, J. and Lee, L.T. (1991) *Langmuir*, **46**, 1855.
53. Kegel, W.K., Bodnar, I. and Lekkerkerker, H.N.W. (1995) *J. Phys. Chem.*, **99**, 3272.
54. Sicoli, F., Langevin, D. and Lee, L.T. (1993) *J. Chem. Phys.*, **99**, 4759.
55. De Gennes, P.G. and Taupin, C. (1982) *J. Phys. Chem.*, **86**, 2294.
56. Daillant, J., Bosio, L., Benattar, J.J. and Meunier, J. (1989) *Europhys. Lett.*, **8**, 453.
57. Shneider, M.B., Jenkins, J.T. and Webb, W.W. (1984) *Biophys. J.*, **45**, 891.
58. Engelhardt, H., Duwe, H.P. and Sackmann, E. (1985) *J. Phys. Lett.*, **46**, 395.
59. Bivas, I., Hanusse, P., Botherel, P., Lalanne, J. and Aguerre-Chariol, O. (1987) *J. Physique*, **48**, 855.
60. Di Meglio, J.M., Dvolaitzky, M. and Taupin, C. (1985) *J. Phys. Chem.*, **89**, 871.
61. Farago, B., Huang, J.S., Richter, D., Safran, S.A. and Milner, S.T. (1990) *Prog. Colloid Polym. Sci.*, **81**, 60.
62. Safran, S.A. (1983) *J. Chem. Phys.*, **78**, 2073.
63. Milner, S.T. and Safran, S.A. (1987) *Phys. Rev. A*, **36**, 4371.
64. Eastoe, J., Sharpe, D., Heenan, R.K. and Egelhaaf, S. (1997) *J. Phys. Chem. B*, **101**, 944.
65. Gradzielski, M. and Langevin, D. (1996) *J. Mol. Struct.*, **383**, 145.
66. Eastoe, J. and Sharpe, D. (1997) *Langmuir*, **13**, 3289.
67. Rock, P.A. (1969) *Chemical Thermodynamics*. MacMillan, London.
68. Olsson, U. and Wennerström, H. (1994) *Adv. Colloid Interface Sci.*, **49**, 113.

## Chapter 6

# Polymers and Polymer Solutions

*Terence Cosgrove*

### 6.1 Introduction

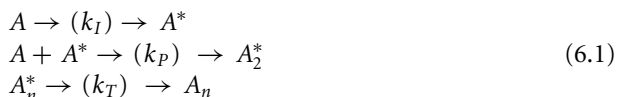
Polymers are long chain molecules that are made by assembling a series of monomers, which may be of the same type (homopolymer) or a mixture (copolymer). Many polymers exist in nature such as natural rubber (*cis*-polyisoprene) and DNA (a polymer with a helical double chain whose backbone is made of an alternating sequence of sugar and phosphates). The number of monomers we can join together is unlimited but stresses in long chains can cause them to break up. Natural polymers can have over  $10^6$  monomers in a chain and all the chains can have the same length. Synthetic polymers commonly have up to  $10^5$  monomers but there is always a distribution of chain lengths. If a typical monomer length is approximately 1 nm then  $10^6$  monomers would have a contour length of 1 mm. Polymers, however, are rarely fully stretched.

In this chapter we review briefly the basic polymerisation schemes for making polymers, the statistics of chain molecules and their solution thermodynamics. More detailed texts in this area include classic texts by Flory [1, 2] and de Gennes [3] as well as several more recent books by Sun [4], Doi [5], Grosberg and Khokhlov [6], Rubinstein and Colby [7] and, for an introduction to polymer synthesis, Stevens [8].

### 6.2 Polymerisation

There are many methods that can be used to make polymers synthetically and these include condensation, free radical and ionic and emulsion methods.

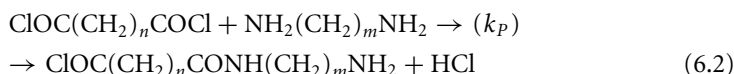
In practice, the final use of a polymer will dictate the most efficient method to choose. Essentially, each of these methods has the same basic chemical steps, but the differences are due to the control of the rate constants associated with them. For example, a fast termination will give only low molecular weight products. The basic steps are initiation (I), propagation (P) and termination (T), each of which can be characterised by a rate constant as in Equation (6.1). In each of the four methods above these rate constants are controlled in different ways and lead to very different molecular weight distributions:





### 6.2.1 Condensation

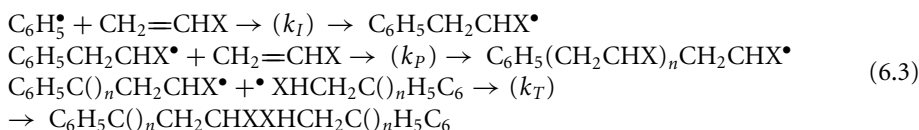
Conceptually (and historically), the method of condensation is the easiest to visualise. In this method two bifunctional monomers are mixed together. The reaction often proceeds spontaneously, as in the famous nylon rope trick, when the ingredients meet at an interface and the product can be pulled out as a thread. Typically, the basic catalysed reaction involves an acid chloride and an amine and, after the first step, both reactive groups are still present in the dimer, Equation (6.2). As the polymerisation progresses both monomers are used and the reaction can proceed until the supply of one or the other is exhausted. The reaction kinetics can be approximated by a second order rate equation and lead to a polydisperse molecular weight distribution:



### 6.2.2 Free radical

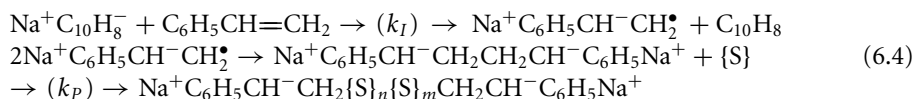
Probably the most common method for polymerisation is the free radical method and an early example was in making poly(vinyl chloride). The initiation step can be carried out using benzoyl peroxide, which attacks the double bond in a vinyl monomer giving a polymeric radical, which can then propagate. The termination can come about by several mechanisms and in the example given in Equation (6.3) a combination route is shown.

The polymers produced by this route are also polydisperse because of the competing reactions (propagation, initiation and termination). Recently, advances in free radical methods have meant that the radical species can be protected and this means that much closer control of polydispersity can be achieved:

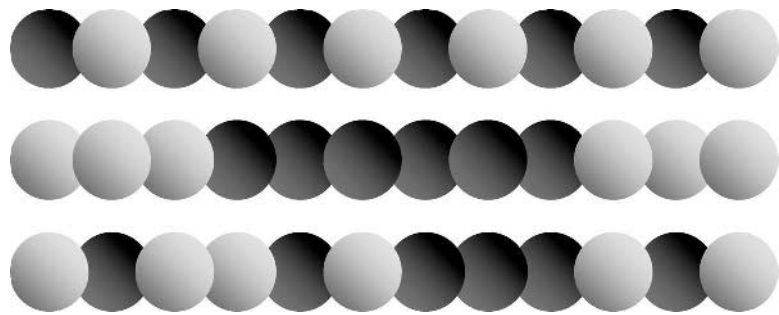


### 6.2.3 Ionic methods

Ionic polymerisation is more technically challenging but is a method for making highly monodisperse polymers, by effectively removing the termination process. If the initiation is very rapid compared to the propagation then living ionic polymers can be made which can be terminated in a controlled way. One example is the polymerisation of styrene (S):



The initiator can be made from sodium and naphthalene in an inert solvent (THF). This gives a carbanion, which combines with styrene to give a radical ion. The dianion can propagate from both ends. The charges stay close because of the low dielectric constant of the medium



**Figure 6.1** The structures of copolymers: the top figure is an alternating copolymer, the middle a block copolymer and the bottom random.

and termination is precluded (provided contamination is kept to a minimum) but can be controlled by quenching. The resulting polymers can be reasonably monodisperse.

6.3 Copolymers

Copolymers come in three different guises: random, alternating and block, as shown in Figure 6.1. Random copolymers can be made by mixing monomers with similar reactivity ratios to give a product, which has the same ratio of monomers as the reactants. They often have properties intermediate between the two monomers used. Block copolymers can be made, in principle, by sequential polymerisation by any of the methods above.

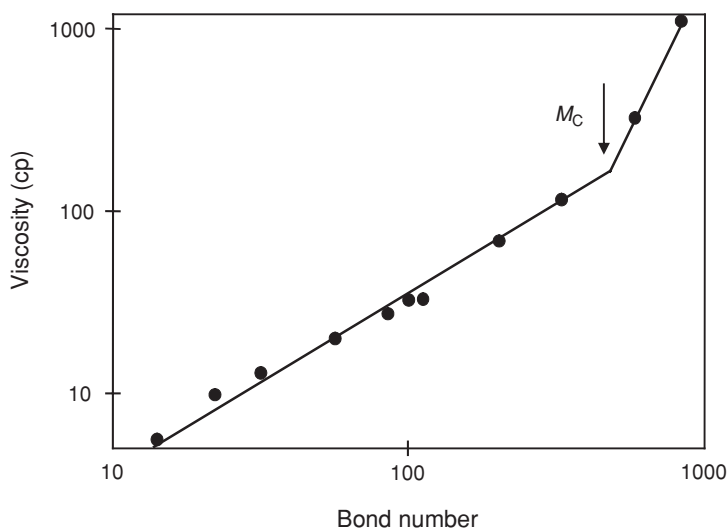
6.4 Polymer Physical Properties

Polymers have very different physical properties to their respective monomers. In Table 6.1 we explore some of these.

One fascinating topological property of polymers is that they can form physical entanglements. This property is strongly molecular weight dependent and gives rise to complex rheological and diffusional properties. For viscosity ( $\eta$ ) the dependence on molecular weight

**Table 6.1** Properties of high and low molecular weight polymers

Property	Low molecular weight	High molecular weight
Solution volume	Little/no change	Small change
Viscosity	Newtonian	Non-Newtonian
Film forming	No/brittle	Yes
Dialysis	No	Yes
Stress	Fractures	Viscoelastic
Ultrasound	No effect	Degradation



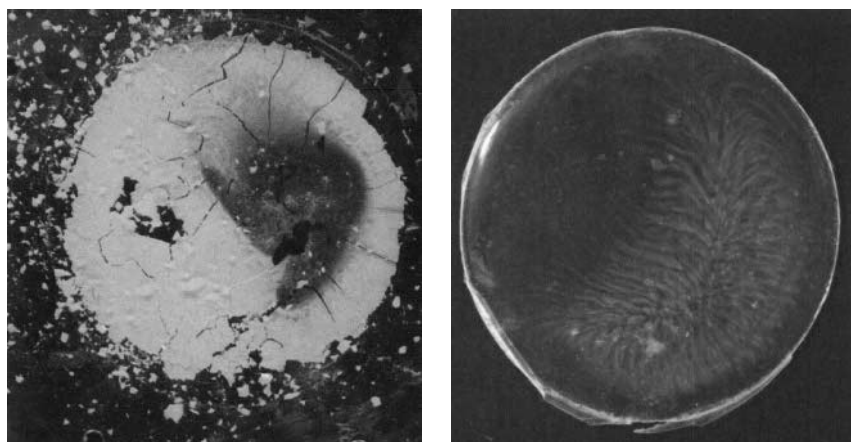
**Figure 6.2** Melt viscosities of polydimethylsiloxane.

switches from one which is linear in molecular weight ( $M$ ) below a critical  $M_c$  to a 3.4 power law above it, as given in Equations (6.5) and (6.6):

$$\eta \sim M \quad (6.5)$$

$$\eta \sim M^{3.4} \quad (6.6)$$

This switchover is seen in many experimental studies and Figure 6.2 shows how viscosity depends on molecular weight for a series of melt polymers of polydimethylsiloxane. Similarly films formed by polystyrene are very sensitive to  $M_c$  as shown in Figure 6.3.



**Figure 6.3** Films formed by evaporating polystyrene solutions of different molecular weights above and below the critical entanglement weight. The figure on the left is a sample of  $20 \text{ kg mol}^{-1}$  and on the right that of  $51 \text{ kg mol}^{-1}$ .

**Table 6.2** Polymer uses

Property	Usage
Impermeability	Protective coatings, beer glasses
Inertness	Artificial joints, prosthetics
Adsorption	Crystallisation modifiers, colloidal stabilisation, adhesives
Strength	Building materials
Electrical	Conducting polymers and insulators
Fluidity	Lubricants, viscosity modifiers

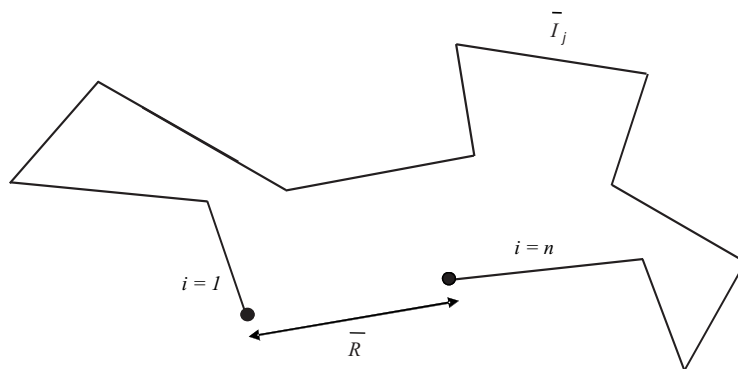
## 6.5 Polymer Uses

Polymers find many uses (Table 6.2) in a wide range of industries from pharmaceutical to heavy engineering.

## 6.6 Theoretical Models of Polymer Structure

Although the contour length of polymer chains may be substantial, because of bond rotation and elasticity, they are rather compact. Flory first suggested that a polymer chain should undergo a random walk through space as shown in Figure 6.4. Ignoring the volume of the chain this model predicts that the end to end distance,  $R$ , is proportional to the square root of the number of bonds,  $n$ .

The proof of the random walk for  $n$  steps is quite straightforward using the vector model. Each step ( $i$ ) is defined by a vector  $\vec{l}_i$ . As the walk has an equal probability of going in any direction, the sum of these vectors, the end to end distance, is zero. We can calculate,

**Figure 6.4** An idealised polymer walk.

however, the mean square end to end distance:

$$\begin{aligned}\langle R^2 \rangle &= \sum_{i=1}^n \vec{l}_i \sum_{j=1}^n \vec{l}_j \\ \langle R^2 \rangle &= \sum_{i \neq j}^n \vec{l}_i \cdot \vec{l}_j + \sum_i^n \vec{l}_i^2\end{aligned}\quad (6.7)$$

The first term in Equation (6.7) is effectively zero as any vector  $\vec{l}_i$  will be at a random orientation to any vector  $\vec{l}_j$  and hence the resultant products will be equally positive or negative and cancel in the summation. The second term is just a scalar, so

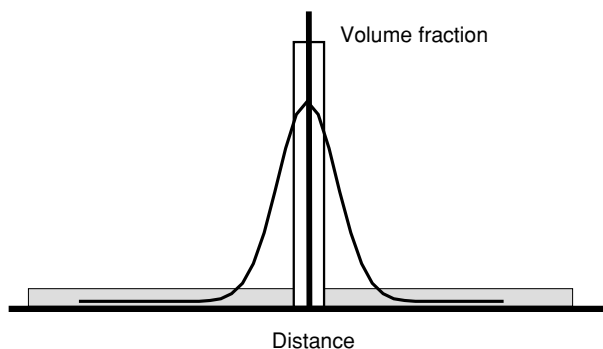
$$\langle R^2 \rangle = n l^2 \quad (6.8)$$

which shows that the end to end distance is proportional to the square root of the number of steps. This is the same result that one gets for a random walk diffusion process.

### 6.6.1 Radius of gyration

It is not straightforward to measure  $R$  directly but several experimental methods allow us to measure the radius of gyration,  $R_G$ , of a polymer chain by, for example, viscosity or scattering. The value of  $R_G$  depends not only on the chain length but also on the shape of the molecule. Equation (6.9) defines  $R_G$ , and  $r_i$  is the distance of monomer  $i$  from the centre of mass. Hence long rods have a very large radius of gyration. Figure 6.5 illustrates how shape determines  $R_G$  for a random coil, a solid sphere and a rod. This means that if we measure  $R_G$  then we must be careful in equating this value to the molecular size if we do not have some indication of the molecular shape:

$$\langle R_G^2 \rangle = \frac{\sum_{i=1}^n m_i r_i^2}{\sum_{i=1}^n m_i} \quad (6.9)$$



**Figure 6.5** The values of the radius of gyration for a polyethylene chain of 50 segments.

**Table 6.3** Values of the characteristic ratio

Polymer	$C_\infty$
Ideal chain	1.0
Poly(ethylene oxide)	5.6
Polydimethylsiloxane	5.2
Polystyrene	9.5
Polyethylene	5.3
DNA	$\sim 600$

For a random walk, Debye showed that  $R_G$  and  $R$  are closely related:

$$\langle R_G^2 \rangle = \frac{\langle R^2 \rangle}{6} \quad (6.10)$$

### 6.6.2 Worm-like chains

Real chains have fixed valence angles and rotations about bonds are not entirely free, so the simplistic result above needs to be modified for real chains. We introduce the characteristic ratio  $C_\infty$ , which can vary between 4 and 20 to correct for chain flexibility. Some examples are given in Table 6.3 which are average values. The solvent here plays an important role (see Section 6.8). We can rewrite (6.8) as:

$$\langle R_G^2 \rangle = \frac{C_\infty n l^2}{6} \quad (6.11)$$

Kuhn showed that any chain could be scaled so that it could be described as completely flexible. This is done by dividing the chain into Kuhn lengths which are multiple numbers of monomers depending on  $C_\infty$ . Another related dimensionless parameter is the persistence defined as,  $p = C_\infty/6$ .

### 6.6.3 Radius of gyration in ideal solution

Rewriting Equation (6.11) in a simpler form allows us to use tabulated values to find the radius of gyration of polymers in ideal solvents (defined as  $R_G \sim n^{0.5}$ ). We use the expression, where  $M$  is in  $\text{gmol}^{-1}$ .

$$\langle R_G \rangle = \alpha M^{0.5} \quad (6.12)$$

and use values of  $\alpha$  in Table 6.4.

### 6.6.4 Excluded volume

For the ideal chains we have discussed above we have used the random walk model, which allows chains to overlap. This is, of course, not very realistic as the chains must have a finite volume to exist. Theoretically this is a tricky area but computer simulations and a speculation

**Table 6.4** Values of  $\alpha$  to calculate the radius of gyration from Equation (6.12)

Polymer	$\alpha/10^{-4}$ nm
Polyethylene	435
Poly(ethylene oxide)	330
Poly dimethylsiloxane	250
Polystyrene	282

by Flory have shown that we can use Equation (6.13) to account for excluded volume:

$$\langle R_G^2 \rangle \sim M^\nu \quad (6.13)$$

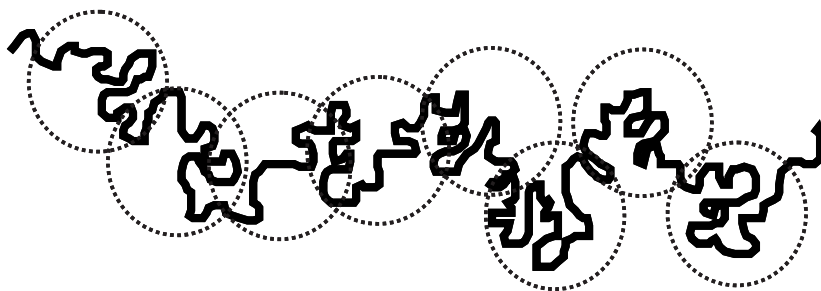
The exponent  $\nu$  is given by  $\nu = 6/(D + 2)$ , where  $D$  is the spatial dimension. So in 3D,  $\nu = 6/5$  and the chain expands beyond its ideal dimensions ( $\nu = 1$ ). The effect does not seem very strong but for high molecular weight chains it is appreciable. Consider a chain with  $10^4$  segments. Then  $R_G$  is predicted by (6.13) to increase by a factor  $\approx 2.5$  over the ideal chain. The ideal chain and the excluded volume chain are two possible models we can use to estimate the polymer chain dimensions. In practice the ideal chain model works in two real situations, in a poor solvent which causes the chain to collapse due to a net chain–chain attraction (Section 6.8) and in a polymer melt where the intra-chain–chain repulsion is balanced by repulsions from neighbouring chains.

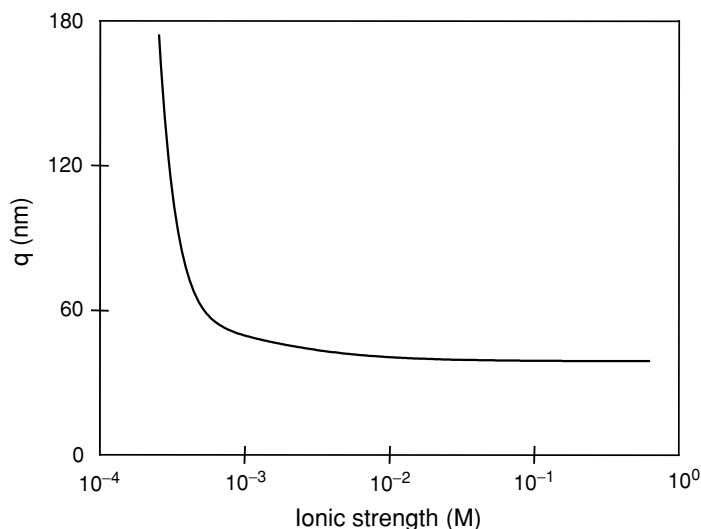
### 6.6.5 Scaling theory: blobs

A rather clever model which encapsulates these two models of polymer chains has been proposed by de Gennes. His idea was to break the chain up into blobs as shown in Figure 6.6. Inside the blob the chain is self-avoiding but the blobs themselves can overlap and are essentially ideal. If we have  $g$  monomers per blob and the blob size is  $\xi$ , then  $\xi \sim g^{\nu/2} \sim g^{3/5}$ . If the blobs can overlap then  $R \sim \xi^{0.5}$  and hence:

$$R \sim n^{0.5} g^{0.1} \quad (6.14)$$

For an ideal chain  $g = 1$  and for a chain with full excluded volume  $n = g$  which recovers the two extreme cases of an ideal and swollen chain above.

**Figure 6.6** The Blob model.



**Figure 6.7** The change in persistence length of DNA in aqueous salt solutions.

### 6.6.6 Polyelectrolytes

Polyelectrolytes are a very common class of polymer in which each monomer carries a charge. For weak polyelectrolytes this is pH dependent (e.g. poly(acrylic acid)) but is pH independent for strong polyacids (e.g. poly(styrene sulphonate)). As the effect of charge is repulsive these molecules can become highly extended and their size will depend strongly on the solution environment (e.g. added salt). Because of this strong stretching we can use the following expression for the end to end distance,

$$\langle R^2 \rangle = L^2(1 - L/3q) \quad (6.15)$$

where  $L$  is the contour length and  $q$  is the electrostatic contribution to the persistence length. Figure 6.7 shows how  $q$  varies with ionic strength for DNA solutions with change in ionic strength.

## 6.7 Measuring Polymer Molecular Weight

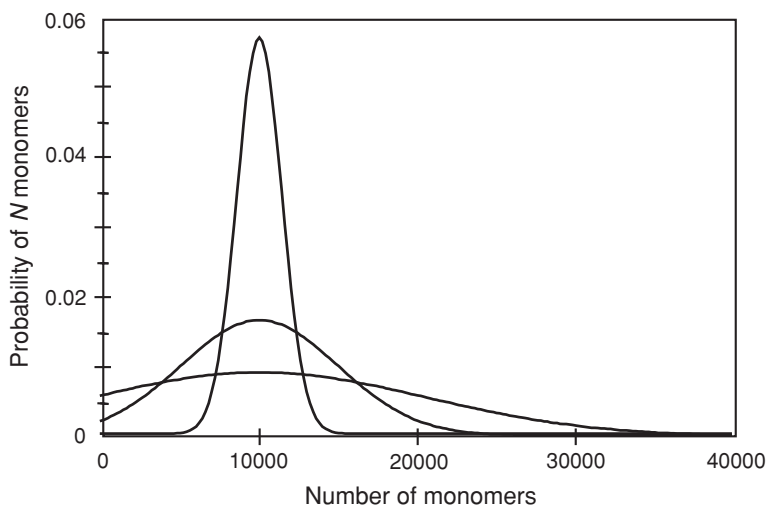
There are many methods to establish the molecular weight of a polymer. However, the problem is not straightforward, as polymer chains are not all of the same length and methods which can distinguish between chain lengths are particularly important. Firstly we shall define what we mean by molecular weight.

The number average molecular weight ( $M_N$ ) is calculated by finding the total weight and dividing by the number of molecules:

$$M_N = \frac{\sum_{i=1}^N N_i M_i}{\sum_{i=1}^N N_i} \quad (6.16)$$

$N_i$  is the number of molecules with mass  $M_i$ .





**Figure 6.8** Molecular weight distribution functions for three different values of  $M_W/M_N$ : 1.38 (broad peak) 1.20 (medium peak) and 1.02 (narrow peak). The value of  $M_N$  was 10,000.

The weight average ( $M_W$ ) is defined as

$$M_W = \frac{\sum_{i=1}^N N_i M_i^2}{\sum_{i=1}^N N_i M_i} \quad (6.17)$$

From these definitions higher molecular weight molecules contribute more to  $M_W$  than to  $M_N$ . For a monodisperse polymer  $M_W$  equals  $M_N$  and the ratio of  $M_W/M_N$  can be used to indicate the degree of polydispersity. Figure 6.8 illustrates the molecular weight distributions for different values of this ratio and it can be seen that even when  $M_W/M_N$  is less than 1.02 there is still appreciable polydispersity.

More generally, higher moments,  $n$ , can be defined as

$$M_n = \frac{\sum_{i=1}^N N_i M_i^n}{\sum_{i=1}^N N_i M_i^{n-1}}$$

Many techniques allow us to measure the molecular weight of polymers and, as they depend on different physical properties, they give rise to different moments of the distribution. Table 6.5 summarises some common methods and the molecular weight average they measure.

A detailed description of polymer viscosity is beyond the scope of this chapter. However, the relationship of viscosity to molecular weight is very important. We can define specific viscosity as a dimensionless quantity with first order solvent effects removed. First defining the specific viscosity  $\eta_{SP}$  as

$$\eta_{SP} = (\eta_{\text{solution}} - \eta_{\text{solvent}})/\eta_{\text{solvent}} \quad (6.18)$$

**Table 6.5** A summary of different polymer molecular weight methods and the moments of the distribution that can be measured

Method	Moment	Polydispersity
Osmotic pressure	$M_N$	No/limited MW range
End group analysis	$M_N$	No/limited MW range
Chromatography	$M_W$	Yes/requires calibration
Scattering	$M_W$	Yes (model required/absolute)
Viscosity	$M_V$	No/requires calibration
MALDI TOF	$M_N$	Absolute

Einstein first noted that the specific viscosity was proportional to volume fraction  $-\eta_{SP} \sim \phi \sim cR^3/M$  as the volume of a polymer coil is  $\sim R^3$ .

The intrinsic viscosity extrapolates this value to infinite dilution to remove intermolecular effects,  $[\eta] = \eta_{SP}/c$  as the concentration  $c$  tends to zero.

As the end to end distance for the polymer coil  $R \sim M^{0.5}$ , we can use this to estimate the coil volume and hence

$$[\eta] \sim R^3/M \sim M^{3/2}/M = kM^a \quad (6.19)$$

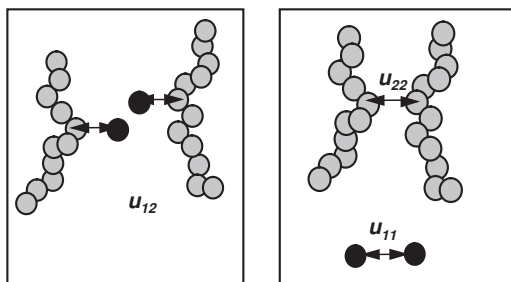
$k$  and  $a$  are known as the Mark–Houwink parameters and the equation gives us a simple semi-empirical formula for the viscosity which we can use to measure molecular weight.

## 6.8 Polymer Solutions

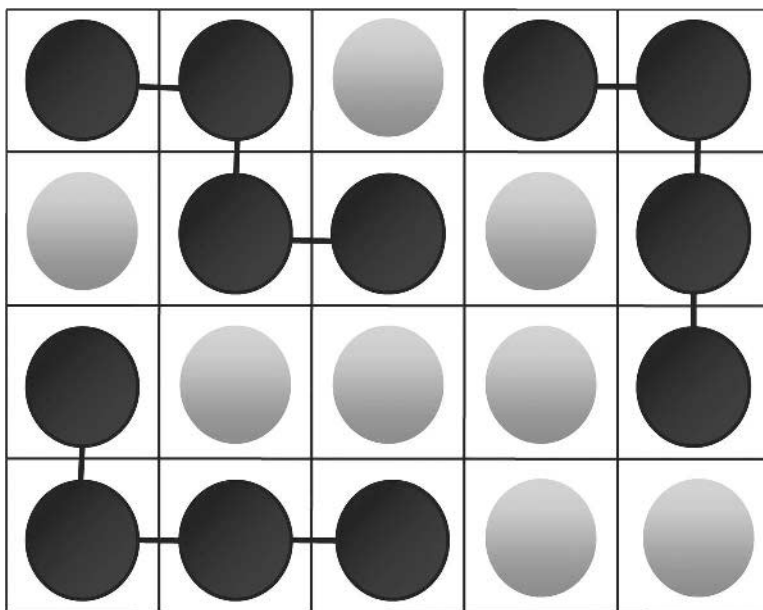
The conformation of polymer chains in solution depends strongly on their architecture and the solvent quality. Flory and Huggins developed a theory for the solubility of polymers based on a net solvent–polymer interaction energy. It is defined as

$$\chi = [(u_{12} - \frac{1}{2}(u_{11} + u_{22}))] \frac{z}{k_B T} = \frac{\Delta u}{k_B T} \quad (6.20)$$

where 1 refers to solvent and 2 to a polymer segment.  $z$  is a coordination number. The final and initial states that lead to this definition are given in Figure 6.9.



**Figure 6.9** The final (left-hand figure) and initial (right-hand figure) states for mixing pure polymer with pure solvent.



**Figure 6.10** The Flory–Huggins lattice.

The Flory–Huggins lattice model (Figure 6.10) can be used to find the free energy  $\Delta A$   $\Delta A_m$  for mixing of a polymer with a solvent and thus to construct a polymer solution phase diagram. The basic model makes the following assumptions.

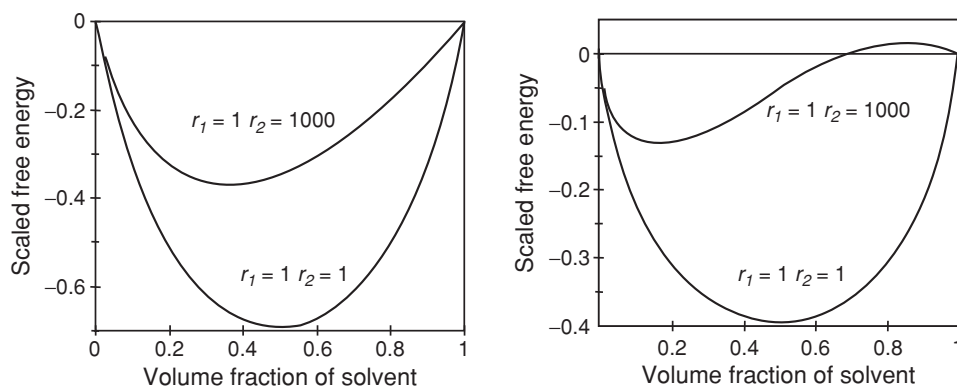
1. The lattice is full with either polymer segments or solvent molecules.
2. The segment size and solvent size are the same.
3. There is random mixing in the lattice.
4. The system is homogeneous and the mean field assumption can be used.
5. The Flory–Huggins parameter is purely enthalpic.

The assumptions make the model reasonably simple to use but they also impose several limitations; most of these can be overcome. The final equation for the free energy,  $\Delta A_m$ , of mixing is given by

$$\Delta A_m = k_B T N \left[ \frac{\phi_1}{r_1} \ln \phi_1 + \frac{\phi_2}{r_2} \ln \phi_2 + \phi_1 \phi_2 \chi \right] \quad (6.21)$$

where  $r$  is the number of monomers/chain ( $r_1 = 1$  for the solvent),  $\phi$  is the volume fraction and  $N$  is the number of molecules (one solvent, two polymers).

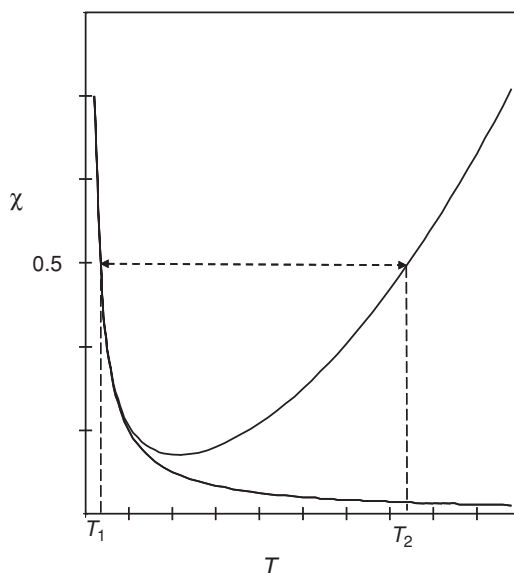
From Equation (6.21) it can be seen that the entropy term is always negative and so dissolution depends on the value of the  $\chi$  parameter. Figure 6.11 shows how the free energy changes with volume fraction of polymer for different values of  $\chi$ . When  $\chi = 0$  we have complete mixing for monomer solution (lower curve) and polymer (upper curve) solutions. When  $\chi = 1.2$ , the monomer solution (lower curve, right-hand figure) is still miscible over the entire volume fraction range, but for the polymer, a miscibility gap appears and the solution breaks down into two immiscible solutions: one dilute and one concentrated.



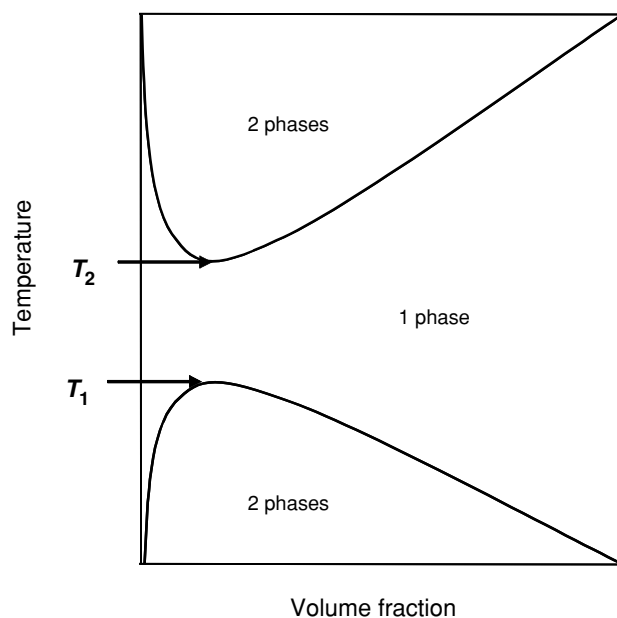
**Figure 6.11** Free energy calculations using the Flory–Huggins model as a function of the  $\chi$  parameter: left-hand figure  $\chi = 0$  and right-hand figure  $\chi = 1.5$ . The upper curves correspond to  $r_1 = 1$  and  $r_2 = 1000$  and the bottom curves correspond to  $r_1 = 1$  and  $r_2 = 1$ .

In terms of a phase diagram, the theory thus predicts an upper critical solution temperature (UCST) and this has been found experimentally for many non-aqueous polymer solutions. However, many aqueous polymer solutions also show a lower critical point as well (LCST). This can be rationalised as a breakdown in the assumption that the system has no free volume. The  $\chi$  parameter can be modified to take this into account

$$\chi = \frac{\Delta u}{kT_B} + AT^n \quad (6.22)$$



**Figure 6.12** Variation in  $\chi$  with temperature using Equation (6.22).



**Figure 6.13** A polymer solution phase diagram. The UCST and LCST are shown as  $T_1$  and  $T_2$ .

where  $T$  is temperature and  $A$  and  $n$  are constants. The form of Equation (6.22) is shown in Figure 6.12 and the resultant phase diagram in Figure 6.13.

A useful experimental parameter that we can obtain directly from Equation (6.21) is the osmotic pressure by obtaining the chemical potential for the solvent:

$$\frac{\Pi}{c} = RT \left( \frac{1}{M} + (0.5 - \chi) \frac{c}{v_o \rho^2} \right) \quad (6.23)$$

This gives another route to determine molecular weight and the  $\chi$  parameter.  $v_o$  is the molar volume of solvent and  $\rho$  is the polymer density. The equation is also the starting point for developing the theory of light scattering from polymer solutions.

Equation (6.23) also reveals another interesting fact that there is a critical value for  $\chi$ . When  $\chi = 0.5$  Equation (6.23) reverts to the ideal van't Hoff equation. Under these conditions the solution is ideal and the temperature at which this occurs is called the  $\theta$  temperature. The balance of the osmotic pressure, which in a poor solvent compresses the chain, with the chain excluded volume, which leads to expansion, accounts for the ideal random coil behaviour of polymer chains that can be observed.

The behaviour of a polymer melt, which also shows ideal behaviour, can also be rationalised in that the osmotic pressure of the polymer in the solutions can also be calculated and shown to be ideal when extrapolated to a pure melt.

Copolymers having more than one monomer type have more than one  $\chi$  parameter. This of course is the basis behind their usefulness as they can disperse in oil or in water as micelles. This aggregation will depend on (among other things) the  $\chi_{AB}$  between the blocks and the block ratio, as well as on temperature and concentration. Many block copolymer systems show very complex phase behaviour. The Flory–Huggins theory can be applied in

principle to these systems to calculate the critical micelle concentrations and other solution properties.

Random copolymers often have intermediate properties between the two monomers, if they are truly random, and it is possible to use a weighted  $\chi$  parameter to describe some of their solution behaviour.

## References

1. Flory, P.J. (1953) *Principles of Polymer Chemistry*. Cornell University Press, Ithaca.
2. Flory, P.J. (1989) *Statistical Mechanics of Chain Molecules*. Hanser, Munich.
3. de Gennes, P.-G. (1979) *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca.
4. Sun, S.F. (1994) *Physical Chemistry of Macromolecules*. Wiley, New York.
5. Doi, M. (1996) *Introduction to Polymer Physics*. Clarendon Press, Oxford.
6. Grosberg, A.Y. and Khokhlov, A.R. (1997) *Giant Molecules*. Academic, San Diego.
7. Rubinstein, M. and Colby, R. (2003) *Polymer Physics*. Oxford University Press, Oxford.
8. Stevens, M.P. (1999) *Polymer Chemistry*. Oxford University Press, Oxford.

# Chapter 7

## Polymers at Interfaces

*Terence Cosgrove*

### 7.1 Introduction

In many situations we need to control the stability of dispersions, for example in pharmaceutical preparations, paints and inks. In other situations we may need to flocculate them, for example in bacterial harvesting or the drying of a paint film. Adsorbed polymers can play a key role in circumventing and controlling these situations. In order to discover how this can be achieved it is necessary to understand the basic structure of an adsorbed polymer and the importance of the solution and surface chemistry.

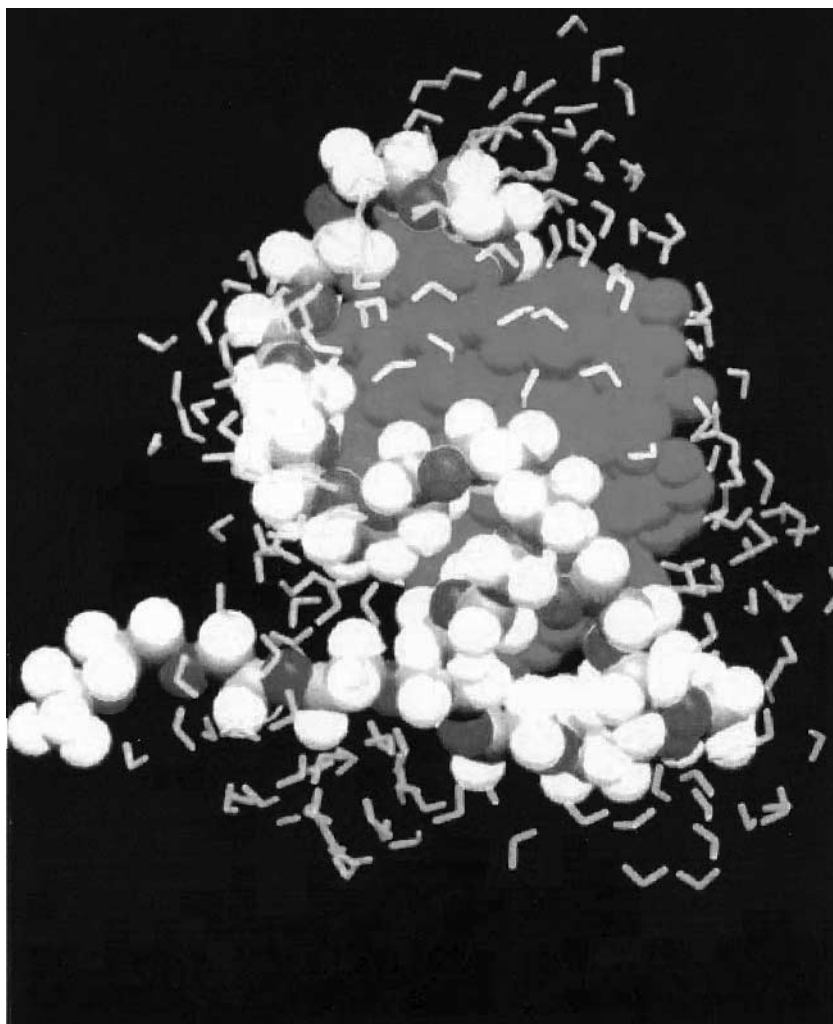
In this chapter we shall introduce some of the basic concepts of polymer adsorption from a theoretical point of view and then make comparisons of the background theory with experiment. The emphasis in this chapter is on discerning parameters of the adsorbed layer which are useful in constructing formulations with desired properties.

Figure 7.1 shows an atomistic simulation of an adsorbed polymer interacting with a nanoparticle of silica in water. In this case the polymer both adheres to the surface and has a long tail which protrudes into the solution. This is the generally accepted picture of an adsorbed polymer with some segments attached to the surface as trains, which are joined by segments in loops. The ends of the chains which often protrude from the surface are tails. It is the balance of these three populations that gives the adsorbed layer its unique properties. Two books which contain detailed accounts of polymer adsorption are those by Fleer *et al.* [1] and Jones and Richards [2]. Several review articles have also been published [3, 4].

#### 7.1.1 Steric stability

Except for some special cases (e.g. microemulsions) colloidal dispersions are not thermodynamically stable but by virtue of energy barriers much greater than  $k_B T$  they are kinetically stable.

Figure 7.2 shows the inter-particle potential between a pair of colloidal particles (details can be found in Chapter 3). This system is kinetically stable by virtue of a surface charge which gives rise to a repulsion (positive) which exceeds the inherent attractive potential (negative) which is due to van der Waals forces. The thermal energy must overcome this repulsion for the two particles to flocculate. However, this balance is very strongly affected by the presence of salt. By increasing the salt concentration the electrostatic repulsion is



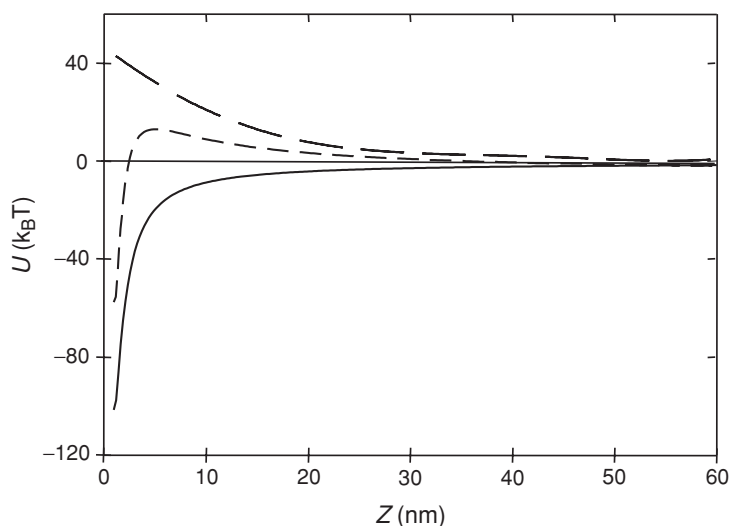
**Figure 7.1** An atomistic simulation of a chain of poly(ethylene oxide) adsorbed onto a silica nanoparticle in water.

easily reduced so that the total potential becomes attractive and the sample will flocculate as shown in Figure 7.3.

### 7.1.2 *The size and shape of polymers in solution*

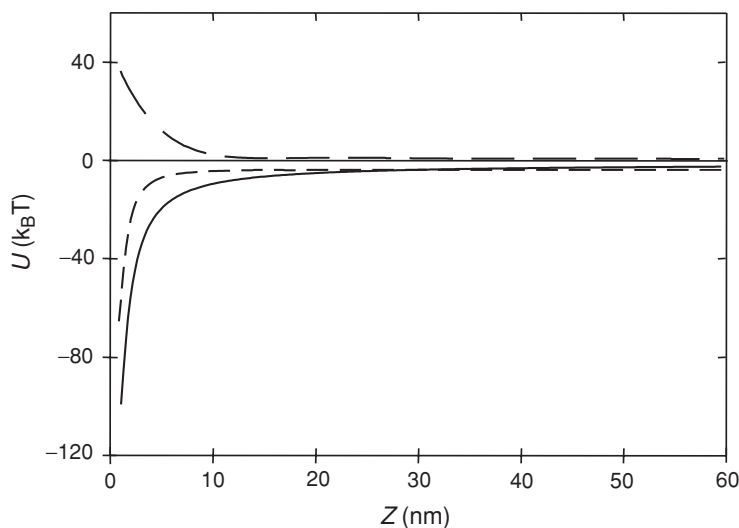
In ideal conditions the radius of gyration,  $R_G$ , of a macromolecule depends on  $N^{0.5}$  where  $N$  is the number of monomers (Chapter 6). This prediction is realised both in a polymer solution and in a polymer melt. In the former case the osmotic pressure of the solvent





**Figure 7.2** The inter-particle potential between two AgBr particles with radius 100 nm in a 0.001 M NaCl solution. (— electrostatic repulsion, --- total potential, — van der Waals attraction).

overcomes the excluded volume of the polymer chain and ideal behaviour is found. This special condition is known as a  $\theta$  solvent and is characterised by the Flory–Huggins parameter  $\chi$  equal to 0.5. The more usual case is when the osmotic pressure does not overcome the excluded volume and the polymer chain expands (a good solvent). In this case  $\chi$  is less than 0.5 and we find that the polymer coil expands leading to  $R_G \sim N^{0.6}$  where  $R_G$  is the radius



**Figure 7.3** The inter-particle potential between two AgBr particles with radius 100 nm in a 0.01 M NaCl solution. (— electrostatic repulsion, --- total potential, — van der Waals attraction).

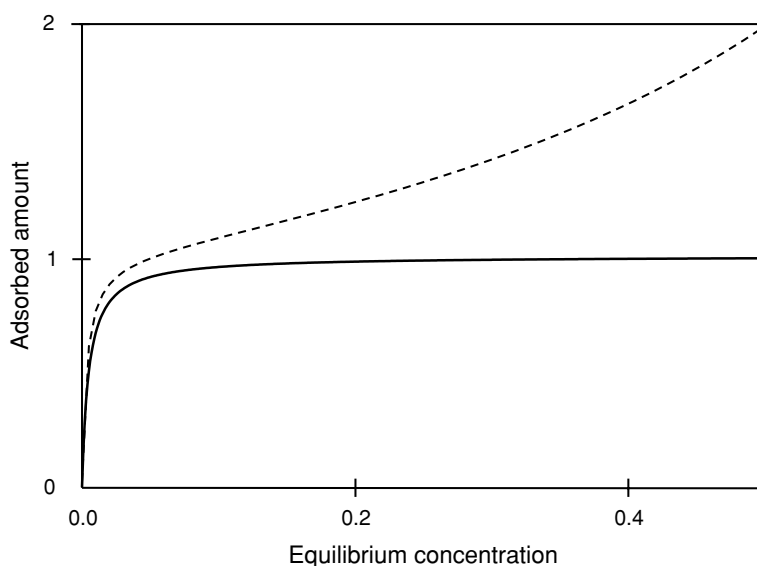
of gyration. The exponent in this relationship is directly related to  $\chi$  (further details can be found in Chapter 6). The shape of an uncharged homopolymer in an ideal solution is approximately spherical, but for block copolymer(s) and polyelectrolytes a whole range of shapes exist and the shape of a molecule strongly influences  $R_G$ .

### 7.1.3 Adsorption of small molecules

It is useful to start by considering the adsorption of small molecules at surfaces. These can be described in many cases by either the Langmuir or the BET isotherm(s) [5]; the former being a limiting case of the latter. Equation (7.1) gives the form of the Langmuir equation where  $\theta$  is the fractional surface coverage,  $c$  is the equilibrium concentration and  $b$  is a constant:

$$\theta = \frac{bc}{1 + bc} \quad (7.1)$$

The model assumes that solute molecules only interact with the surface and, hence, only a monolayer can be formed. Figure 7.4 illustrates the form of the equation. The BET model includes solute–solute interactions and, hence, a multi-layer structure can be formed, as is evident in Figure 7.4, for a particular set of interaction energies. For polymer adsorption there are examples in the literature of using the Langmuir equation to describe experimental isotherms, but, although functionally the equation often works, interpretation of thermodynamic variables such as the adsorption energy must be treated with some caution. The reason for this is the effect of polydispersity (see Section 7.4.2).



**Figure 7.4** A comparison of the Langmuir (solid Line) and BET (dashed line) isotherms.

## 7.2 Adsorption of Polymers

### 7.2.1 Configurational entropy

Unlike small molecules there is often a large configurational entropy penalty to pay when a polymer adsorbs from solution on to a surface. We can use a simple thermodynamic argument to estimate the change in entropy when a polymer adsorbs. For simplicity we assume that each polymer segment has 3 possible orientations ( $\Omega$ ). That means that for  $N$  segments a polymer has  $3^N$  possible conformations. Similarly, if the coil adsorbs completely flat, then in 2 dimensions there are  $2^N$  conformations.

We can estimate the entropy change from the third law of thermodynamics as

$$S = k_B \ln(\Omega) \quad (7.2)$$

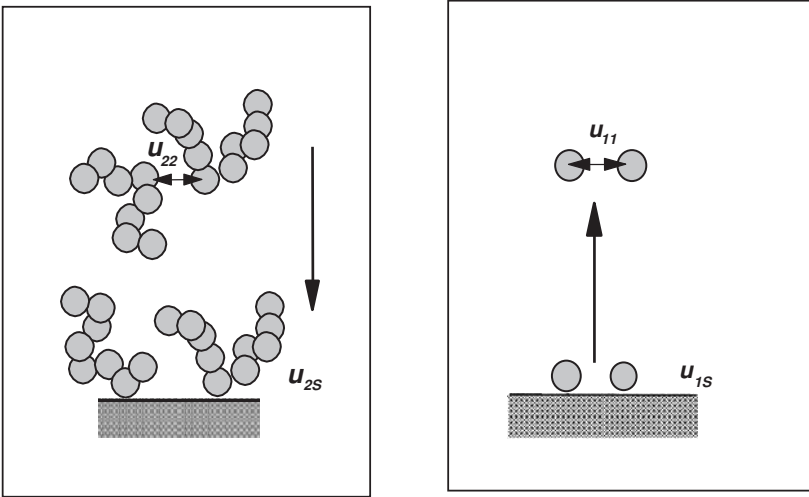
Hence

$$\Delta S = k_B \ln(2^N/3^N) \text{ or } \sim -0.4k_B \text{ per monomer} \quad (7.3)$$

For this entropy penalty to be overcome we need a critical enthalpy of at least  $\sim 0.4k_B T$ . This simple analysis ignores the effect of liberating the solvent from the surface which is clearly entropically advantageous.

### 7.2.2 The Flory surface parameter $\chi_s$

In a similar way to the definition of the Flory–Huggins solution parameter,  $\chi$ , we can define a Flory surface parameter  $\chi_s$  which is defined in terms of an initial state comprising a polymer melt and adsorbed solvent and a final state of bulk solvent and adsorbed polymer as shown in Figure 7.5. The reason for this particular form of this definition is that it makes



**Figure 7.5** The exchange process of adsorbing a polymer segment and displacing an adsorbed solvent molecule (1) refers to the solvent, (2) to a polymer segment and  $a$  refers to adsorbed.

$\chi$  and  $\chi_s$  effectively independent i.e. we separate the surface interactions from the solution interactions. In reality this would be very difficult to achieve.

The definition of  $\chi_s$  is given by

$$\chi_s = - \left[ (u_{2s} + u_{1s}) + \frac{1}{2}(u_{11} - u_{22}) \right] / k_B T \quad (7.4)$$

where  $u$  are the pairwise contact energies and are normally attractive ( $<0$ ).  $S$  refers to the surface, 1 a solvent molecule and 2 a polymer segment. For adsorption to take place  $\chi_s$  must be greater than the critical value as given by Equation (7.3).

## 7.3 Models and Simulations for Terminally Attached Chains

Because many aspects of the adsorption of polymers depend on chain conformations, simulation methods, which explore the different ways in which a polymer can interact with an interface, are particularly revealing. The simplest example is to constrain the polymer chain to a lattice rather than free space and to attach one end of the chain irreversibly to the surface. These two constraints reduce the number of chain conformations possible by a very large amount and prevent the chain from desorbing into the solution.

The simplest of these models involves counting the number of chain conformations exactly and this is the first approach we shall use. For longer chains the approximate methods of Monte Carlo and molecular dynamics are useful both for single and multiple chains. Finally, we shall describe a full thermodynamic model of polymer adsorption where the constraint of terminal attachment is overcome and a full equilibrium of the adsorbed population of chains with those in solution can be achieved.

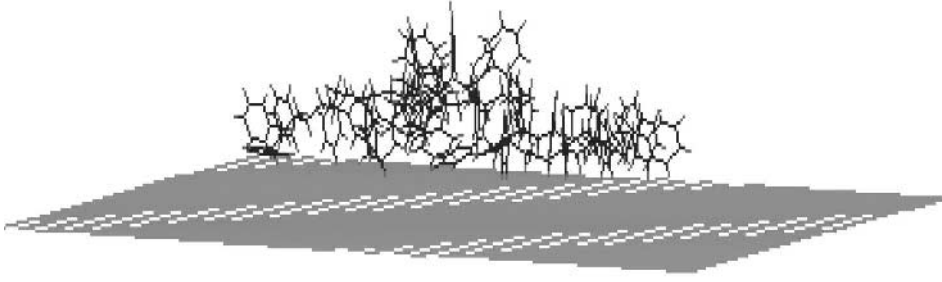
### 7.3.1 Atomistic modelling

Ideally, to predict the interaction of a polymer with a substrate, we would like to build a theory which retains the detailed atomic structures of the solvent, polymer and interface. For simple visualisations this can be done but to make a detailed study at this level for practical systems with an ensemble of long chain lengths is not currently feasible. Nevertheless, the approach is useful and in Figure 7.6 we show a simulation of polystyrene on graphite. One interesting aspect of this visualisation is that the phenyl rings do not lie flat on the surface because of steric constraints. In this case the definition of the bound surface layer has to be generalised to segments within a certain distance of the surface. In these situations it is the shape of the molecular, i.e. the chemistry, which is the determining factor.

In a more realistic model we need to average over all possible chain conformations and the result would be a volume fraction profile normal to the surface which describes the number of segments in layers starting at the interface and finishing in the bulk solution.

Given an explicit form for the volume fraction profile we can readily calculate other important parameters using the following equations:

$$\Gamma = \rho_2 \int_0^{\text{span}} \phi_{\text{ads}}(z) dz \quad (7.5)$$



**Figure 7.6** An atomistic level simulation of a single polystyrene chain on a graphite surface. Segments in trains (touching the surface), loops and tails can be seen.

$$p = \rho_2 \int_0^{\text{bound layer}} \phi_{\text{ads}}(z) dz / \Gamma \quad (7.6)$$

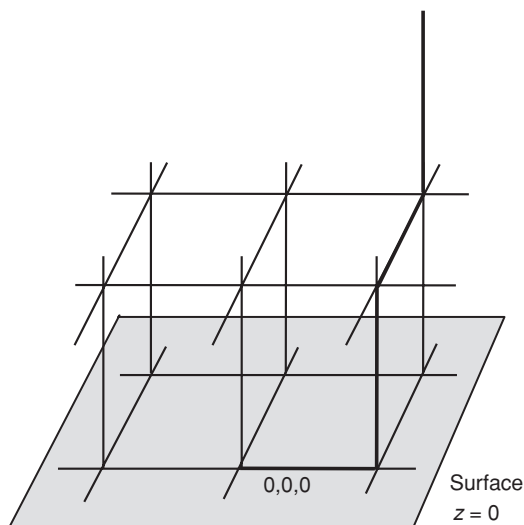
$$\delta_{\text{RMS}}^2 = \rho_2 \int_0^{\text{span}} \phi_{\text{ads}}(z) z^2 dz / \Gamma \quad (7.7)$$

The adsorbed amount  $\Gamma$  is just the integral under the profile multiplied by  $\rho_2$  the polymer density, giving units of mass/unit area. The bound fraction  $p$  corresponds to fraction of segments in the first adsorbed layer and is typically a monomer length in width. The span in this context is the point at which the profile is zero (the maximum value is the chain length). In order to make a comparison of the experimental variables as above with a lattice model, a procedure for scaling to real space is required. There are several strategies for this, for example each lattice site could be occupied by one monomer or one statistical segment.

### 7.3.2 Exact enumeration: terminally attached chains

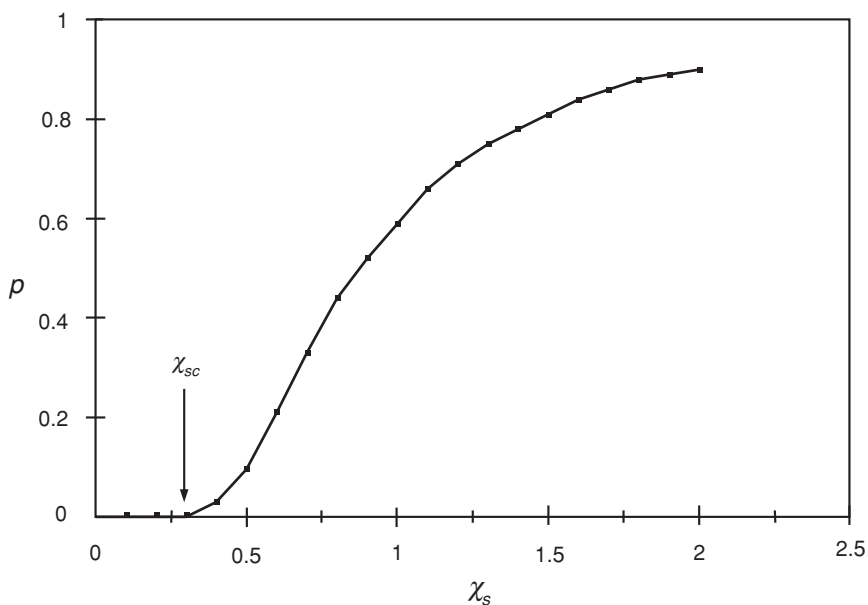
In its simplest form this approach uses a lattice with a single chain terminally attached to the interface. The basic idea is just to count every possible conformation: exact enumeration (EE). For example we shall use a cubic lattice as shown in Figure 7.7. We start at the coordinates (0, 0, 0) and reject all conformations which penetrate the surface and violate the excluded volume criterion (i.e. occupy the same lattice point).  $C(N, m)$  is defined as the number of walks of length  $N$  bonds of which  $m$  are in the layer next to the interface. So, for example,  $C(1, 1) = 4$  and  $C(2, 1) = 4$ . This soon becomes very difficult, and to reach more than  $N = 20$  takes an appreciable amount of computer time; even this modest chain length can have over  $10^9$  conformations. Once we have enumerated the array of numbers  $C(N, m)$  we can quite easily calculate the bound fraction,  $p$ , of the walk. This is the fraction of the total number of segments that is in the layer next to the surface i.e. as trains. For example, for the walk shown in Figure 7.7  $p(N) = 1/4$ . The results for all the conformations are combined using a Boltzmann weighting term which is just  $e^{m\chi_s/k_B T}$ , and Equation (7.8) shows the statistical sum needed to find the average value of  $p$ ,  $\langle p \rangle$ :

$$\langle p \rangle = \frac{\sum_{m=1}^N C(N, m) m e^{\frac{m\chi_s}{k_B T}}}{N \sum_{m=1}^N C(N, m) e^{\frac{m\chi_s}{k_B T}}} \quad (7.8)$$

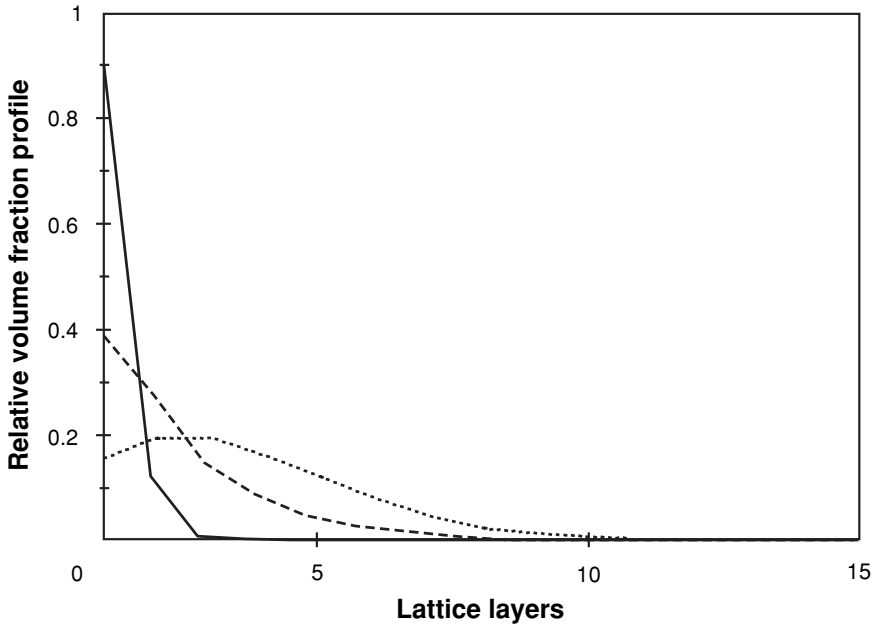


**Figure 7.7** A typical self-avoiding walk on a cubic lattice.

It is useful to explore how  $\langle p \rangle$  depends on the net adsorption energy  $\chi_s$  but the calculation so far is for a finite value of  $N$  and the chain is anchored irreversibly at the surface. This means that the limiting value of  $\langle p \rangle$  for  $\chi_s = 0$  is  $1/N$ . The model can be developed further by using an extrapolation method to find the value of  $\langle p \rangle$  in the limit of large  $N$  and these data are shown in Figure 7.8. At values of  $\chi_s$  less than the critical value of  $\chi_{sc}$  the bound



**Figure 7.8** The variation in the bound fraction  $\langle p \rangle$  as a function of the adsorption energy  $\chi_s$ .



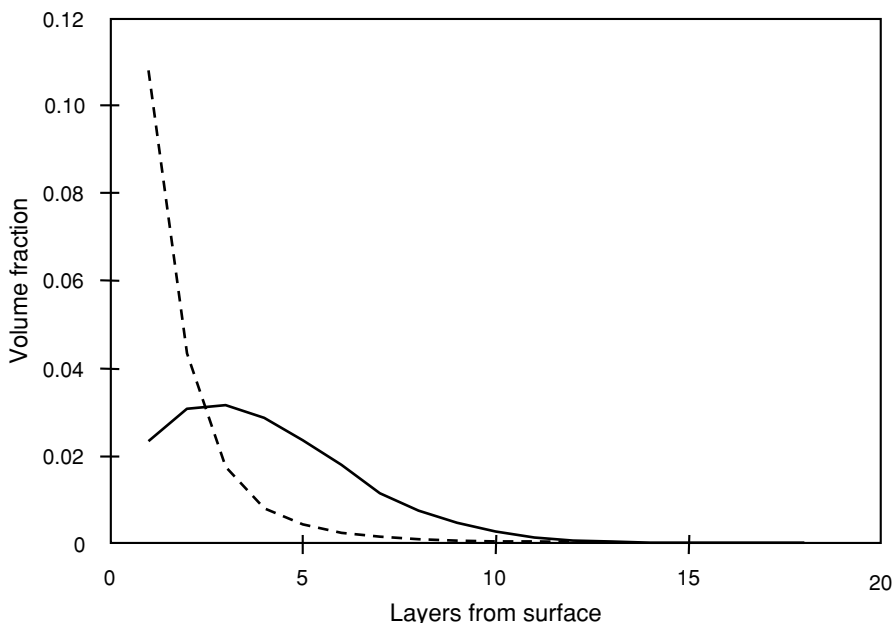
**Figure 7.9** The proportion of segments in layers normal to the surface for a terminally attached chain of 15 segments on a cubic lattice as a function of the adsorption energy  $\chi_s$ . Three values are shown 0.0 (.....), 0.6 (-----) and 1.8 (—).

fraction is zero which would correspond to no adsorption. However, below the critical value there is adsorption. In the limit as  $\chi_s$  becomes much greater than  $k_B T$  for an isolated chain all the segments tend to lie in the first lattice layer as trains.

This simple approach can be easily extended to work out the average number of segments in each of the lattice planes parallel to the surface. This requires some extra counting as we need to know the number of walks  $C(N, m, s, z)$  which have  $m$  surface contacts and  $s$  segments in layer  $z$ . The data in Figure 7.9 show how the shape of the relative volume fraction profile varies as we change  $\chi_s$ . At the surface we see the same picture as above. Below the critical adsorption energy, the number of segments in the first layer decreases. In this case we are dealing with a finite chain length ( $N = 15$ ) and as these conformations all have one segment irreversibly attached at the surface,  $p$  is greater than 0. This regime has become known as a *mushroom* since  $\phi(z)$  has a maximum. In contrast, when  $\chi_s$  is 1.8, the chain collapses on the surface and this is known as a *pancake*.

### 7.3.3 Approximate methods: terminally attached chains

The approach above is very limited as it is just a single short chain. Another approach is to use Monte Carlo (MC) or molecular dynamics (MD) methods [6]. In these methods not all the chain conformations are generated, but a subset. By using appropriate selection criteria, this subset can be representative of the whole. The methods are not restricted to lattices and can deal with multiple chains of lengths substantially greater than with exact enumeration.



**Figure 7.10** An MC simulation of a terminally attached chain with 50 segments and a surface coverage of 0.15, for two values of the adsorption energy above and below the critical value.

A simple procedure for multiple chains is to use a periodic boundary such that any chain crossing the boundary is re-entered in the opposite side of the cell. This, however, does impose coherence in the chain structures, but is still a useful procedure.

Figure 7.10 shows a chain length of 50 and a surface coverage,  $\theta$ , of 0.15 which is defined as the number of adsorbed segments per surface lattice site. This approach also reproduces the profiles found above, the pancake and the mushroom, but now the volume fraction scale is absolute.

In the molecular dynamics approach, Newton's laws of motion are used to generate new conformations from old ones by evaluating the interaction forces in the system. Typically, simulations can be performed on the nanosecond timescale.

### 7.3.4 Scaling models for terminally attached chains

Another approach to discovering the structure of a terminally attached chain is through the scaling approach of de Gennes [7]. In this model the chain is decomposed into  $g$  blobs as in Chapter 6 and as shown in Figure 7.11.

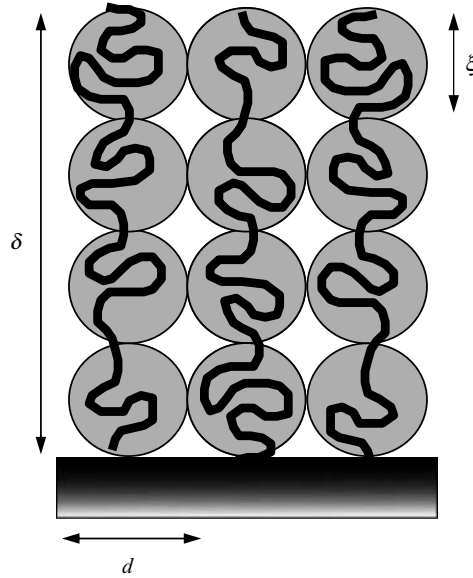
If each blob contains a self-avoiding walk of  $g$  monomers then the blob size  $\xi$  is given by

$$\xi = g^{3/5} \quad (7.9)$$

For  $N$  monomers the brush length,  $\delta$ , is then

$$\delta = \left( \frac{N}{g} \right) \xi \quad (7.10)$$





**Figure 7.11** The blob representation of a series of terminally attached chains.

The second, and central, assumption is that the blob size is directly related to the grafted amount  $\sigma$  so that  $\sigma = 1/\xi^2$ . Combining this result with Equations (7.9) and (7.10) we arrive at:

$$\delta = N\sigma^{1/3} \quad (7.11)$$

This is a very surprising result as it predicts that the brush length is linear in chain length. This is clearly true for a rod normal to the surface but suggests that chains closely grafted together on a surface are very strongly stretched. A more sophisticated approach which confirms this result predicts that the brush volume fraction profile is parabolic [8].

### 7.3.5 Physically adsorbed chains: Scheutjens and Fleer theory

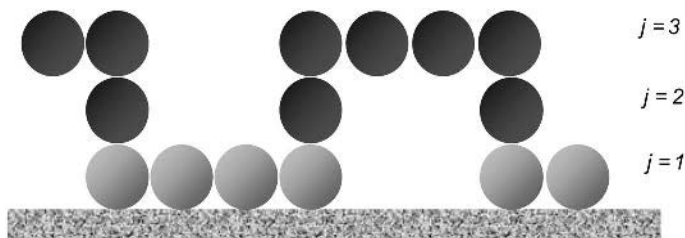
The most successful and useful theory for polymer adsorption is that developed by Scheutjens and Fleer (SF) [1]. This uses the basic ideas of the Flory–Huggins theory for polymer solutions and applies this to each layer of lattice built on a solid substrate. In the SF model, layers, each containing  $L$  sites parallel to the surface, are numbered  $j = 1, 2, 3$ .

The goal of the model is to calculate and minimise the free energy of the system. This can be done in two stages: first calculating the energy and then the entropy.

In the example in Figure 7.12 a single chain conformation is shown which spans three layers. To find the energy  $U$  for this example we just need to find the total number of nearest neighbours for each segment. At the surface, for a cubic lattice we have 6 surface-polymer contacts and 19 polymer-solvent contacts. For three layers,  $j = 1, 2, 3$  we can express the energy  $U$  as

$$U = 6\chi_s + [19\chi + 12\chi + 25\chi]/z \quad (7.12)$$

where  $z$  is the lattice coordination number.



**Figure 7.12** A chain of 15 segments, six of which are at the interface.

This is a useful exercise but a more general approach is required and this can be done using the mean-field approximation; instead of using the actual number of contacts we use the probability that there is a polymer-solvent or polymer surface contact in a given layer. This is then just the volume fraction of polymer in layer  $j$ ,  $\phi^a(j)$ . For example, the average number of contacts in the first layer between the polymer and the surface is just  $L\phi^a(1)$ . More generally, Equation (7.12) can be written as:

$$\frac{\Delta U}{k_B T L} = \phi^a(1)\chi_s + \sum_{j=1}^M \phi^a(j) \langle \lambda \rangle \chi \quad (7.13)$$

The total number of layers is  $M$  and  $\lambda$  is a parameter which corrects for the different number of contacts in the plane and between planes. For example in a cubic lattice there are four nearest neighbours in the plane and one above and one below the plane and these must not be over-counted.

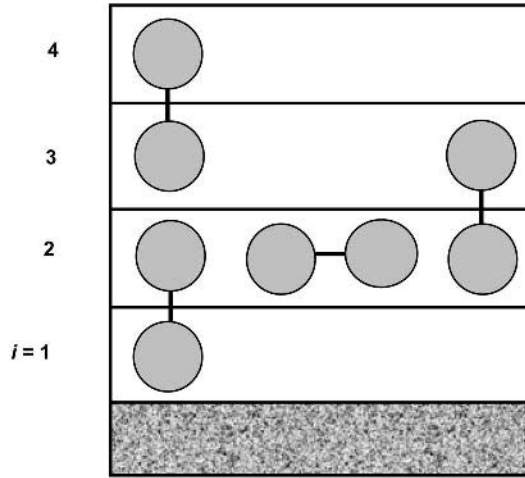
The calculation of the entropy is more involved and a formalism is needed that uses the mean field approach to make up explicit chains. The idea behind this is to use the concept of a free segment weighting factor  $G$ .

$G(z)$  is defined as the *weighting* in the ensemble that a monomer in layer  $z$  has compared to a segment in the bulk of the solution. For non-interacting segments if  $j < 1$  then  $G(z) = 0$  as no monomers can penetrate the surface. Similarly if  $j > 1$  then  $G(z) = 1$  as beyond layer 1 all segments are in the bulk solution (i.e. not interacting with the short range forces at the surface). If  $z = 1$  however then  $G(z) = e^{\chi_s}$  and a segment has a higher probability of being in the surface than in the bulk if  $\chi_s > 0$ .

The next step is to combine these weighting factors to form chains. Starting with the shortest chains possible, which are dimers, we can work out the weighting factors as follows. In Figure 7.13 there are four dimers in different conformations. The dimers are labelled by the lattice layer in which they exist. For example dimer (3, 4) has segment (1) in layer 3 and segment (2) in layer 4. The weighting factor for (3, 4) is given by:

$$G(3, 4) = \frac{G(3)G(4)}{6} = \frac{1}{6} \quad (7.14)$$

The combined weighting factor is arrived at as follows. Each segment on the lattice, not next to the surface (i.e. in layer 1) has a weighting factor of unity. The number of nearest neighbours on the lattice is six and there is only one possible way to construct the dimer (3, 4) starting at segment (3) and so the probability is  $1/6$ .



**Figure 7.13** Four dimers on a cubic lattice in different orientations.

Similarly, dimer (1, 2) has a weighting factor:

$$G(1, 2) = \frac{G(1)G(2)}{6} = \frac{1}{6}e^{\chi_s}$$

Though segment (1) has an extra contribution as it is in layer (1) next to the surface.

The next step is to find an expression for the end-segment weighting factors, which is the weighting for a chain whose end segment is in layer  $j$ . This is effectively the number of chains which start in layers  $(j - 1)$ ,  $(j)$  and  $(j + 1)$  and end in layer  $j$ , and is given by

$$G(j; s) = G(j) \left[ \frac{1}{6}G(j - 1) + \frac{4}{6}G(j) + \frac{1}{6}G(j + 1) \right]$$

where  $s$  is the end segment.

We can use this construction to work out the overall weighting for any chain segment in layer  $j$  by combining two chains whose end segments are in the same lattice layer. So the contribution to the volume fraction in layer  $j$  from any segment  $s$  for a chain with  $N$  segments is

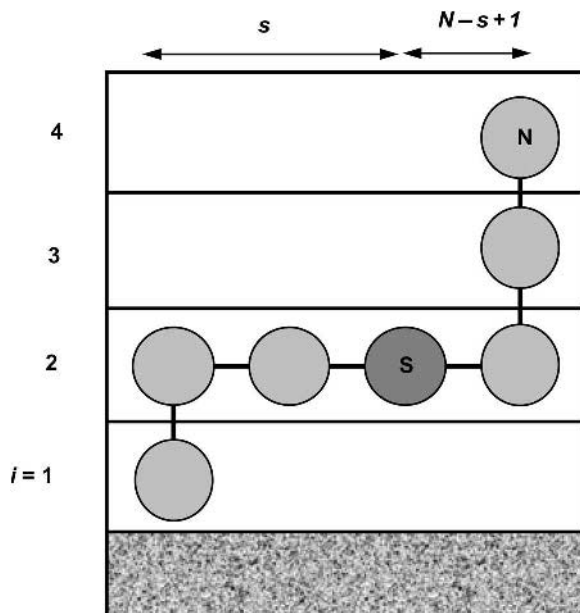
$$\phi(j, s) = \frac{C}{G(j)} G(j; s) G(j, N - s + 1)$$

This is shown schematically in Figure 7.14.

Hence, the total contribution to the volume fraction in layer  $j$  can be found by summing over all segments in the chain:

$$\phi(j) = \sum_{s=1}^N \phi(j; s)$$

This formalism can be extended to calculate explicit volume fraction profiles in equilibrium with a polymer solution. This makes it possible to estimate all the parameters needed to



**Figure 7.14** The connectivity rule for end segments forming a chain of  $N$  monomers.

characterise the adsorbed layer – the adsorbed amount, the layer thickness and the bound fraction – and to make comparisons with experiment.

The volume fraction profiles in Figure 7.15 were calculated for a chain of 50 segments with an ideal solvent (the Flory parameter,  $\chi = 0.5$ ). Two cases are evident:

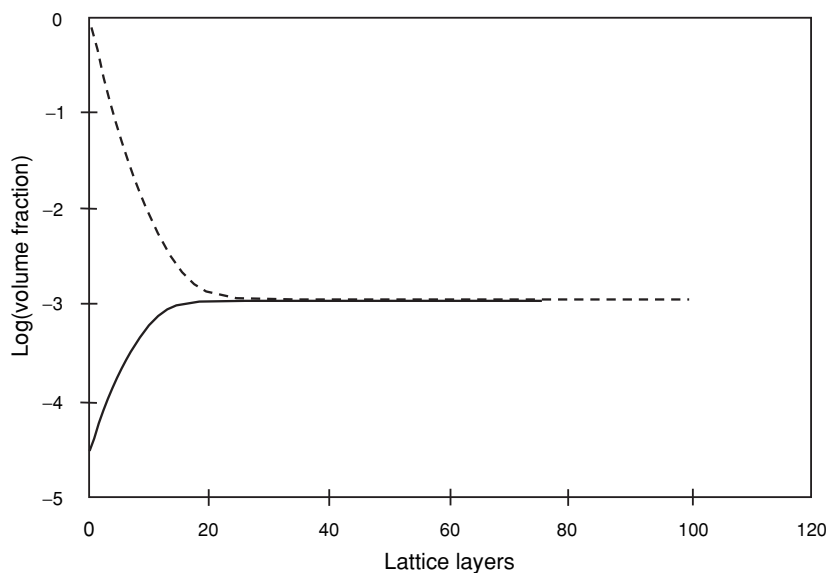
1. when the surface Flory parameter,  $\chi_s$ , is greater than the critical value  $\chi_{sc}$  we have adsorption and the profile falls monotonically to the bulk solution concentration (1000 ppm);
2. when  $\chi_s < \chi_{sc}$  we have depletion and the concentration of segments at the surface is less than the bulk concentration.

Unlike the volume fraction profiles for terminally attached chains (Figure 7.10) chains can be completely desorbed.

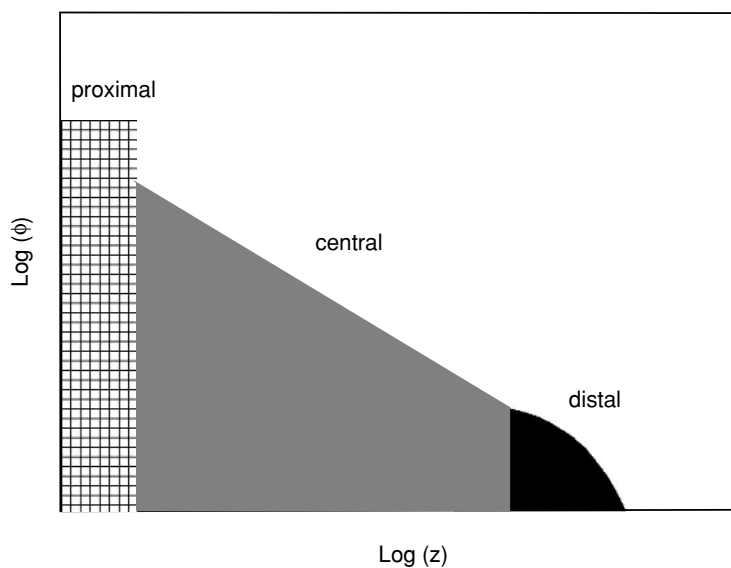
### 7.3.6 Scaling theory for physical adsorption

The scaling approach discussed in Section 7.3.4 can also be used to model a physically adsorbed chain. The basic approach is to treat the adsorbed layer as ‘self-similar’ to a polymer solution. Three regimes are envisaged as shown in Figure 7.16.

- (a) The proximal region which is effectively the train layer, with a width of the order of the monomer length  $a$  i.e.  $\phi \sim \text{constant}$  for  $z \leq a$ .



**Figure 7.15** Volume fraction profiles calculated using the SF model for a chain of 50 segments with two different values of  $\chi_s$ :  $\chi_s < \chi_{sc}$  (solid line) and  $\chi_s > \chi_{sc}$  (dashed line). A more detailed account of this work can be found in the book by Fleer *et al.* [1].



**Figure 7.16** The scaling model of a physically adsorbed polymer layer showing the three regions: proximal (cross-hatched), central (light grey) and distal (black).

- (b) The central region which is similar to a semi-dilute polymer solution. The local volume fraction of a polymer in a semi-dilute solution is given by:

$$\phi \sim N/R^3$$

and in a good solvent  $R \sim N^{3/5}$  (Chapter 6) then:

$$\phi \sim R^{5/3}/R^3 \sim R^{-4/3}$$

The volume fraction at a distance  $z$  from the wall can be thought of as the space in which the chains exist and hence:

$$\phi \sim z^{-4/3} \quad \text{for } a < z < d \quad (7.15)$$

- (c) The distal region which is the periphery of the layer is approximated as an exponential decay:

$$\phi \sim e^{-z} \quad \text{for } d < z < \text{span} \quad (7.16)$$

## 7.4 Experimental Aspects

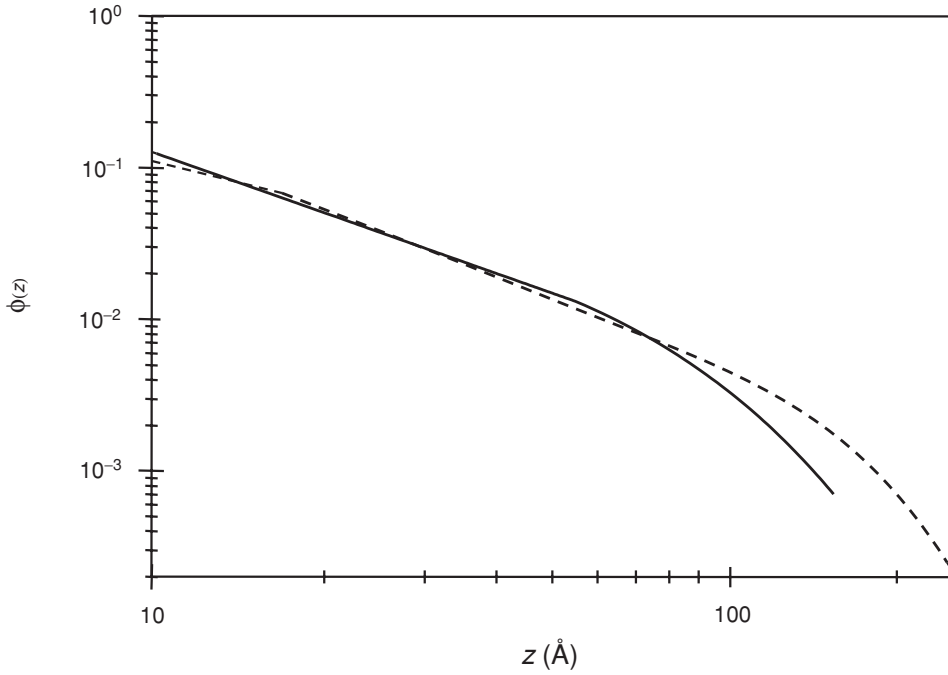
### 7.4.1 Volume fraction profiles

There are several approaches to obtaining experimental volume fraction profiles but for particles the most successful has been to use small-angle neutron scattering (SANS) and for macroscopic surfaces, neutron reflection. One approach is to fit the theoretical profile shapes directly to the scattering data and Figure 7.17 shows an example of the profiles found after fitting such data [9]. The system is poly(ethylene oxide) (PEO) 110K molecular weight adsorbed on polystyrene (PS) latex of radius  $\sim 600$  Å.

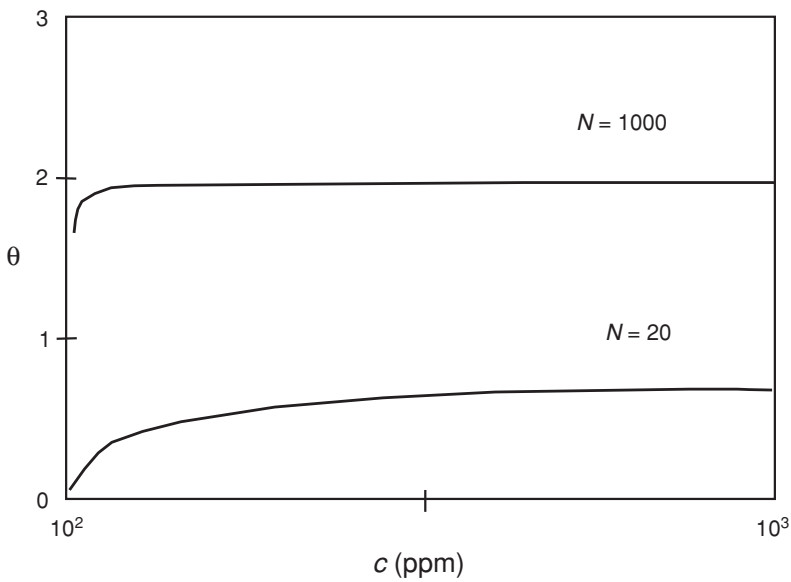
The two different profile shapes are very similar up to  $\sim 80$  Å. Beyond this the tail region is significant and this is calculated explicitly by the SF approach. The sensitivity of the method is  $\sim 0.001$  volume fraction and so for longer chains it underestimates the chain span.

### 7.4.2 Adsorption isotherms

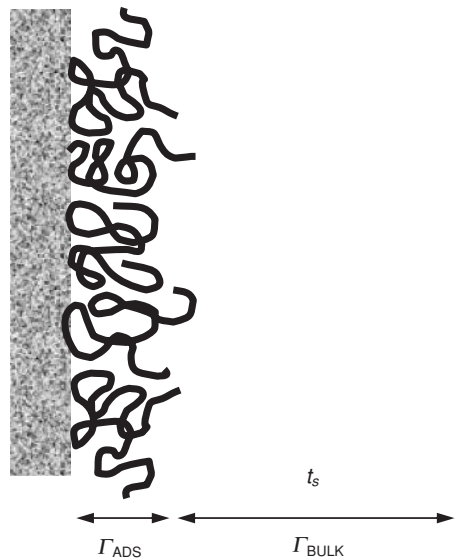
Adsorption isotherms can be calculated from the SF model but not directly from scaling theory. A typical example is given in Figure 7.18 for two different chain lengths. The shorter chain has a lower affinity isotherm, similar to the Langmuir case. The longer chain has a much higher affinity which is typical for monodisperse polymers and shows that below saturation virtually all the added polymer is removed from solution. Under good solvent conditions the plateau level is proportional to  $\log(N)$ . From above it is clear that adsorption is molecular weight dependent and for a polydisperse polymer there are further complications. Entropically, it is more favourable for the longer chains to adsorb though dynamically the short ones may reach the interface first. This competition means that at equilibrium shorter chains will be displaced by longer ones and that dilution may not desorb these chains. The effects can be treated with the SF model and Figure 7.19 illustrates the important parameters. The ratio of the volume of the solution to the available surface area



**Figure 7.17** Experimental volume fraction profiles for PEO 110K adsorbed on polystyrene latex in water, obtained by fitting SANS data to SF (---) and scaling profiles (—).



**Figure 7.18** Theoretical isotherms using the SF model for two chain lengths ( $N = 100$  lower curve and  $N = 1000$  upper curve) with values of  $\chi_s$  of 1.0 and  $\chi$  of 0.5.  $\theta$  is the number of adsorbed segments per surface site.



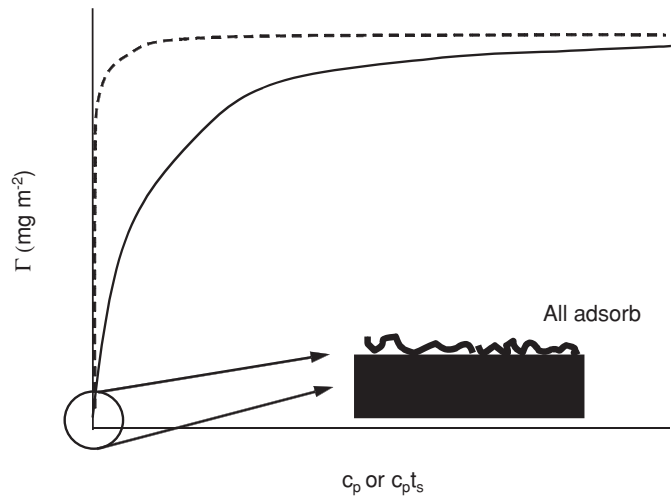
**Figure 7.19** The relation between the surface area and the volume of the solution.

is the most important parameter. In dispersions with very large particles or single flat surfaces these effects can be very pronounced.

The total amount of polymer in the system can be written as

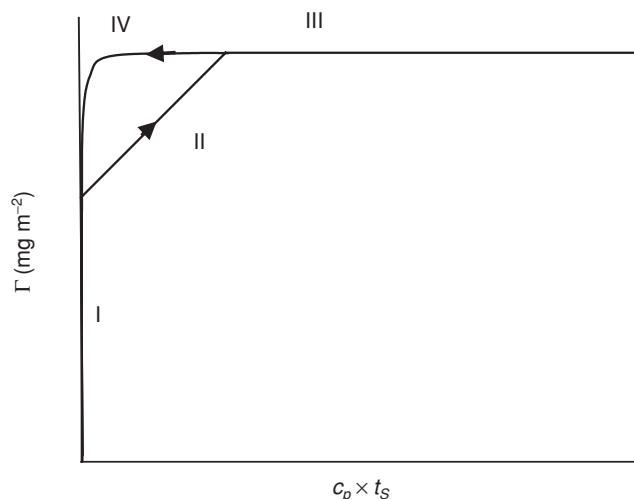
$$\Gamma_{\text{total}} = \Gamma_{\text{ads}} + \Gamma_{\text{bulk}} \quad \text{and} \quad \Gamma_{\text{bulk}} = c_p t_s \text{ [mg m}^{-2}\text{]}$$

This is shown schematically in Figure 7.20.



**Figure 7.20** Schematic adsorption isotherms for two different molecular weight polymers.





**Figure 7.21** Adsorption and desorption from a bimodal mixture of polymer molecular weights.

The effects of polydispersity can be seen clearly by taking just a bimodal polymer solution. The adsorption isotherms for such a situation are shown in Figure 7.21, and where the adsorbed amount is plotted against  $c_p t_s$ , the polymer concentration times the thickness of the solution i.e. the total volume of polymer solution per unit area of surface. As before, at low polymer solution concentrations, virtually all the polymer is adsorbed i.e. both chain lengths.

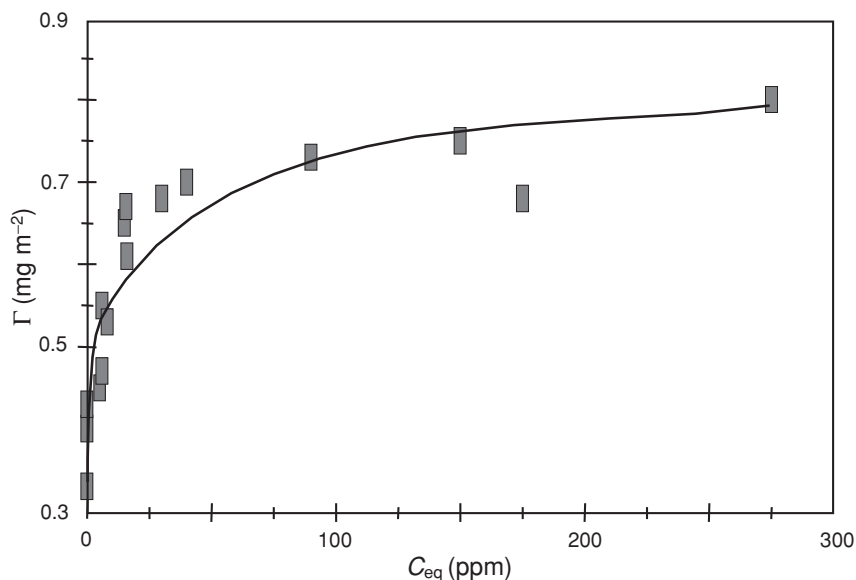
This is region I in Figure 7.21. However, as the solution concentration is increased and we approach saturation (region II) the longer chains are preferentially adsorbed and the small ones that have adsorbed are displaced into the solution. In region III, the surface is saturated. A very interesting and not entirely obvious effect of this adsorption scenario is what happens if we now dilute the polymer solution. This is shown as region IV where the longer chains are left at the surface and complete desorption is now very difficult. The equivalent polymer solution concentrations necessary to achieve this adsorbed state are very low. In some senses this can be seen as an irreversible adsorption. More details can be found in [1].

Another aspect of polydispersity is that the adsorption isotherms become rounded and this means that any attempt to get adsorption energies from the initial slope (by a Langmuir analysis) will be fruitless.

A typical high affinity experimental adsorption isotherm is shown in Figure 7.22. The data shown are for a narrow molecular weight PEO 51K molecular weight adsorbed on PS latex. The isotherm breaks away from the vertical axis at about  $0.4 \text{ mg m}^{-2}$  which corresponds to the onset of inter-polymer interactions in the layer. Values of order  $1 \text{ mg m}^{-2}$  are very typical for the adsorbed amount for uncharged homopolymers.

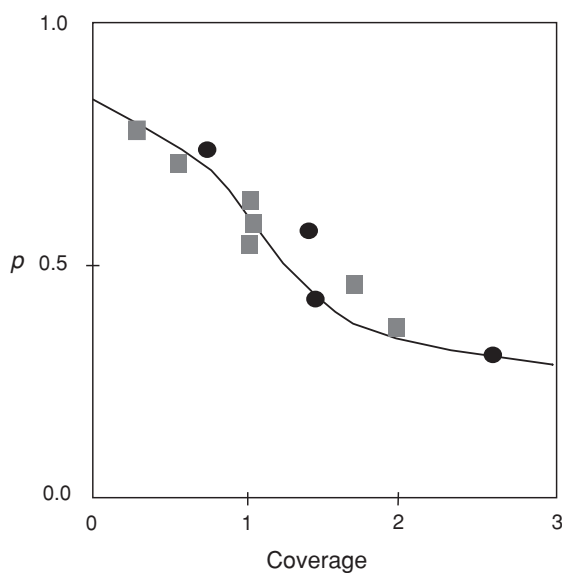
### 7.4.3 The bound fraction

The bound fraction describes how much of the polymer is anchored at the interface. Single chains lie flat on the surface, but for multiple chains this is not so. To optimise the free energy,

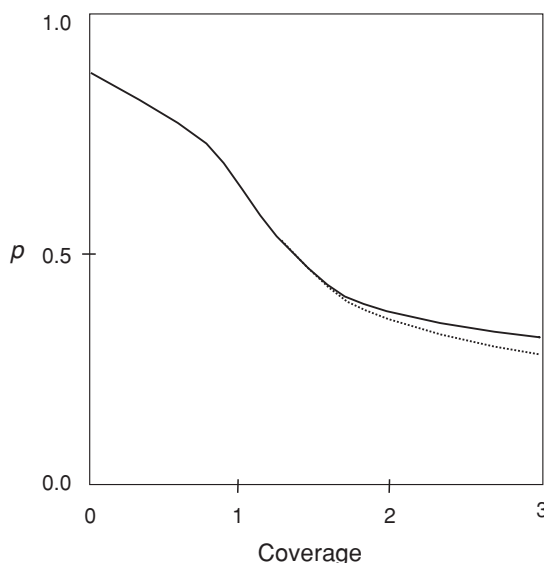


**Figure 7.22** A typical high affinity adsorption isotherm for PEO 51K g mol<sup>-1</sup> adsorbed on to polystyrene latex from water.

surface sites are filled to gain energy but the chains retain a more three dimensional structure at the surface to retain entropy. The experimental data in Figure 7.23 are for poly(vinyl pyrrolidone) (PVP) adsorbed on silica from water and the results have been obtained using two very similar experimental methods, nuclear magnetic resonance (NMR) and electron



**Figure 7.23** Values of the bound fraction,  $p$ , for PVP adsorbed on silica from water. NMR data [●] ESR data [■].



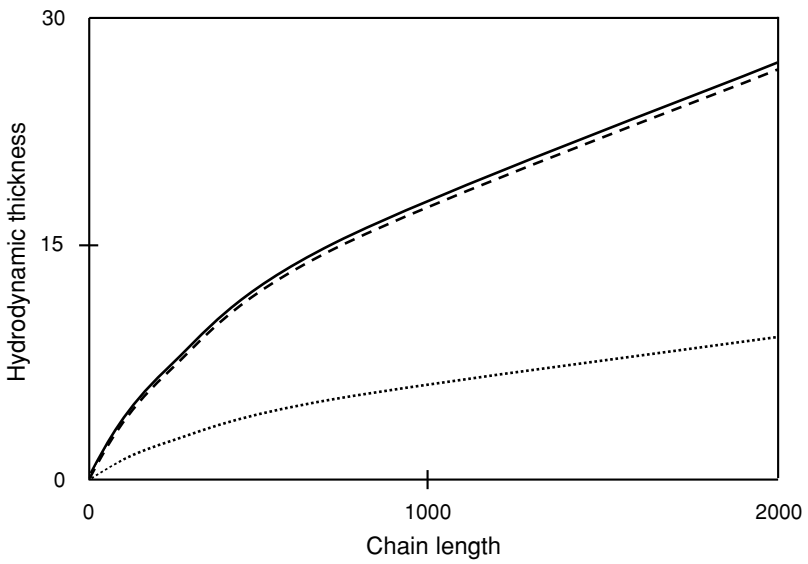
**Figure 7.24** SCF calculations of the bound fraction of an adsorbed polymer as a function of the adsorbed amount for two different chain lengths [100, solid line] and [1000, dotted line].

spin resonance. These methods work on the basis that segments at the interface (trains) are more restricted in mobility than those in loops and tails. Several methods can be used to estimate the bound fraction and these include FTIR, microcalorimetry, solvent NMR relaxation and SANS. For comparison Figure 7.24 illustrates a theoretical calculation of the bound fraction using the SF model. Two different chain lengths have been used. Both theory and experiment show that at low coverages the chains lie relatively flat at the surface (a completely flat conformation would have  $p = 1$ ). With increasing coverage, although the volume fraction at the surface layer will remain fairly constant to incorporate more chains, the adsorb layer will swell, leading to the formation of loops and tails. For a flat layer the bound fraction is independent of chain length, and  $p$  values for different molecular weights will only diverge when there are sufficient number of segments in loops and tails. The SCF theory shows this latter point very nicely, and both theory and experiment confirm this picture of the adsorbed layer.

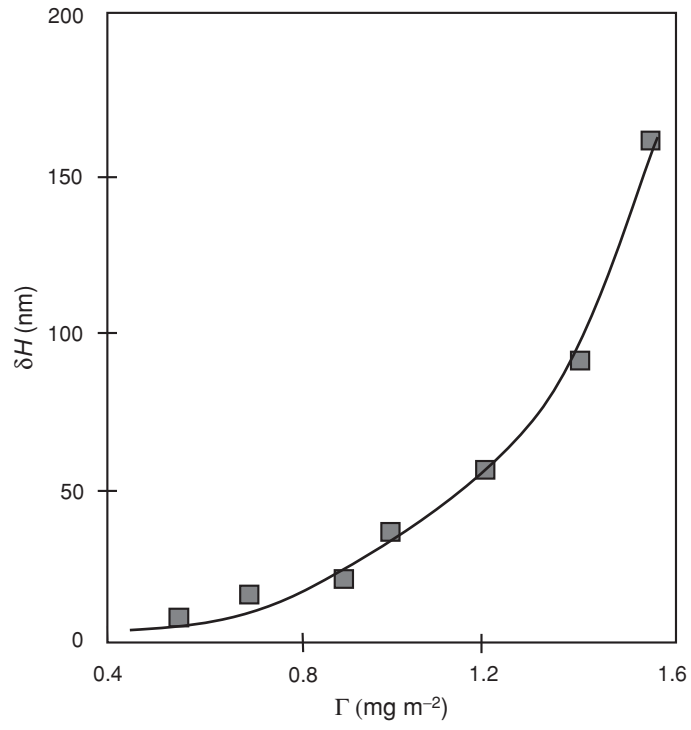
#### 7.4.4 The layer thickness

The layer thickness is a very important parameter for an adsorbed layer as it helps us to design steric stabilisers. Figure 7.25 shows the results of an SCF calculation of the contribution of loops and tails to the hydrodynamic layer thickness,  $\delta_H$ . The contribution of trains can be ignored, but the surprising factor is that tails dominate  $\delta_H$  and become increasingly important as the chain length increases. The tail region is the first point of contact when two particles come together and when the repulsive/attractive steric force is felt.

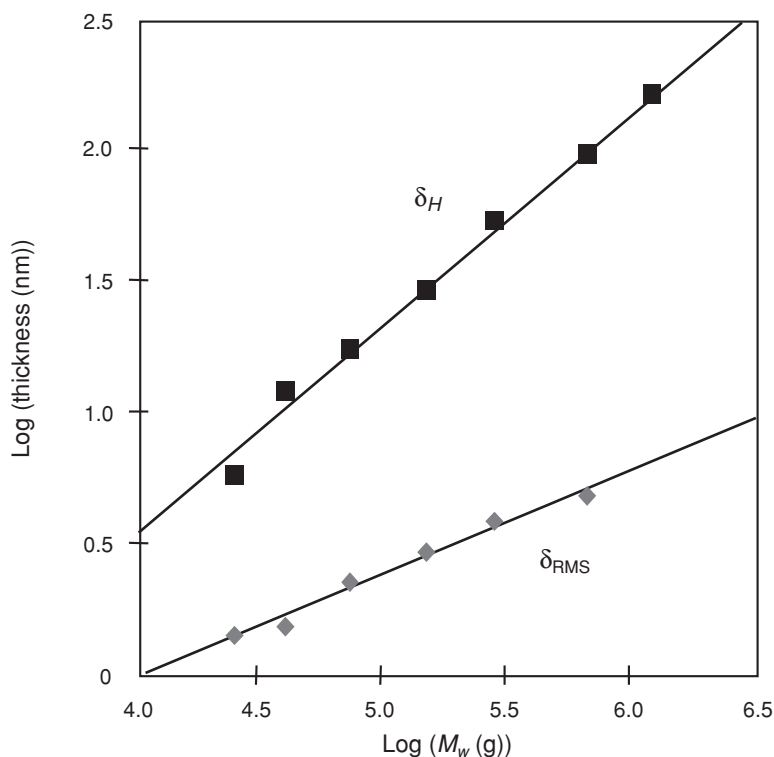
Experimentally, dynamic light scattering (PCS) and viscosity can be used to measure  $\delta_H$  and Figure 7.26 shows a set of data for PEO adsorbed on polystyrene latex as a function of adsorbed amount. The data collapse on a single curve and illustrate the increasing



**Figure 7.25** A theoretical prediction of the contribution of loops (...), tails (- -) and the full profile (—) to the hydrodynamic thickness of an adsorbed layer as a function of chain length.



**Figure 7.26** Hydrodynamic layer thickness for PEO adsorbed on polystyrene latex as a function of the adsorbed amount.



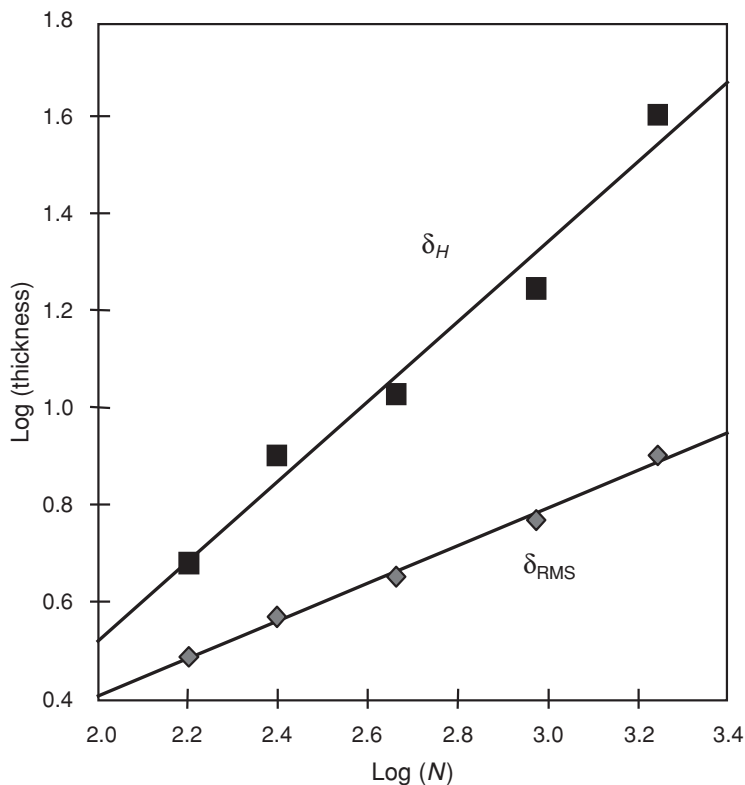
**Figure 7.27** The variation in the hydrodynamic thickness of layers of PEO adsorbed on polystyrene as a function of molecular weight latex measured by photon correlation spectroscopy ( $\delta_H$ ) and SANS ( $\delta_{RMS}$ ).

importance of tails. This behaviour can be understood in that the maximum layer thickness is determined by a combination of the adsorbed amount and molecular weight. For a given molecular weight if there is a maximum in adsorbed amount (the plateau region) then that will be reflected in the maximum attainable layer thickness.

The explicit molecular weight dependence of the layer thickness can be measured by taking a series of samples from the plateau region of the isotherms. Figure 7.27 shows a set of data for the same series of polymers. The data clearly follow a scaling prediction of the form:

$$\delta_H \sim N^a \quad (7.17)$$

The exponent obtained from these data was  $\sim 0.8$  but other workers have reported lower values of around 0.6. For comparison, in the same figure data for the RMS (root mean square) layer thickness are also shown. These data are easily calculated from the volume fraction profile as obtained by SANS [1] or by ellipsometry. The data also show a scaling law but this time the exponent is  $\sim 0.4$ . The hydrodynamic thickness is determined by the tail region as seen above, whereas the RMS thickness is dominated by the central region of the profile, which is mainly loops. Figure 7.28 shows a theoretical calculation of these two thickness parameters based on the SF theory and a percolation model for the hydrodynamics.

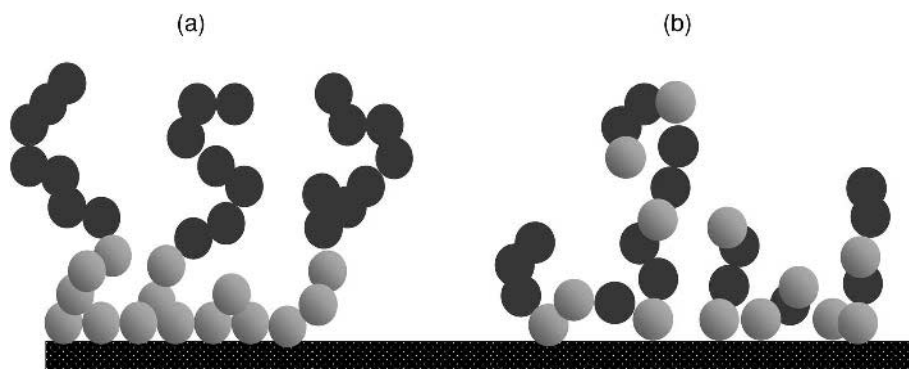


**Figure 7.28** The variation in the hydrodynamic thickness as a function of chain length calculated using the SCF model and a hydrodynamic model.

The agreement between the SF predictions and experiment is good and shows that the hydrodynamic layer thickness increases more strongly with  $N$  than does the RMS.

## 7.5 Copolymers

An adsorbed homopolymer is often a compromise as a stabiliser as it needs to be strongly adsorbed *and* strongly solvated. These two conditions are not compatible as strong solvation implies a small  $\chi$  value which reduces the adsorption preference. To optimise both these effects the system of choice is a copolymer in which one segment type can be strongly adsorbing and poorly solvated and the other the reverse. For a block copolymer there is a strong partitioning at the interface and the adsorbing block can even form a polymer melt at the surface. Figure 7.29 illustrates this partitioning behaviour. The block copolymer forms a train region which is populated mainly by the more strongly adsorbing and less solvated block and a brush region composed of the block with the opposite properties. On the RHS of this figure we also illustrate the adsorption of a random copolymer. The exact surface structure that results depends on the relative ratios of the blocks and their distribution in the chain, but typically we might expect that the copolymer has an intermediate behaviour between

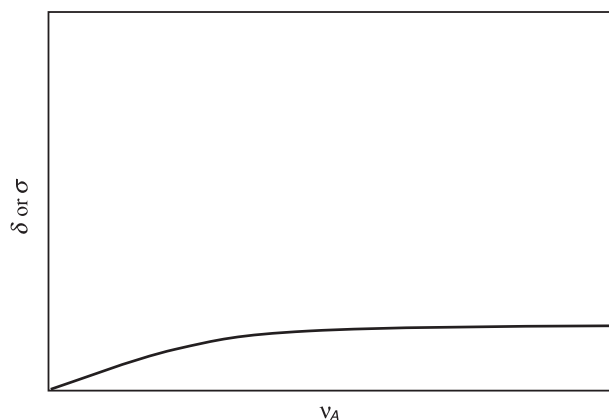


**Figure 7.29** A schematic representation of the adsorption of a block copolymer (LHS) and a random copolymer (RHS). The adsorbing segments are in grey and the non-adsorbing ones in black.

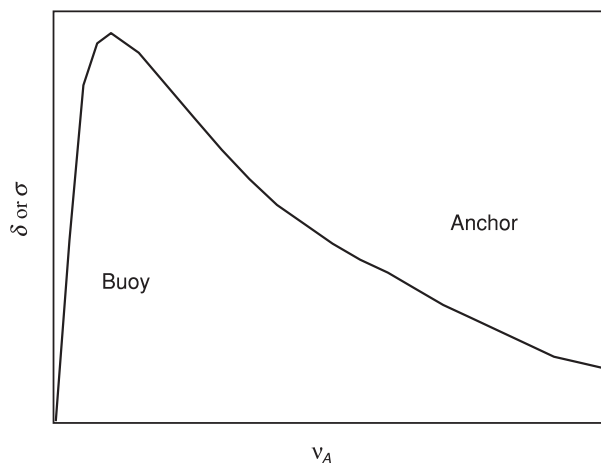
two homopolymers made up of the respective blocks. This is shown very clearly in the SF calculation in Figure 7.30 where one constituent homopolymer does not adsorb and the other does. Both the adsorbed amount and the layer thickness behave in a very similar manner.

For a block copolymer the behaviour is quite different as is seen in Figure 7.31. Here there is a balance between the energy gained from adsorbing the anchor block [A] and the entropy associated with stretching the buoy block [B]. This leads to a maximum in the adsorbed amount and the layer thickness as a function of the fraction of adsorbing segments ( $\nu_A$ ) when these two effects balance. So in choosing a block copolymer there is an optimum block ratio at approximately  $\nu_A \sim 0.2$  to get the maximum steric effect. For a random copolymer the adsorption varies from depletion when  $\nu_A = 0$  to that of homopolymer A when  $\nu_A = 1$ .

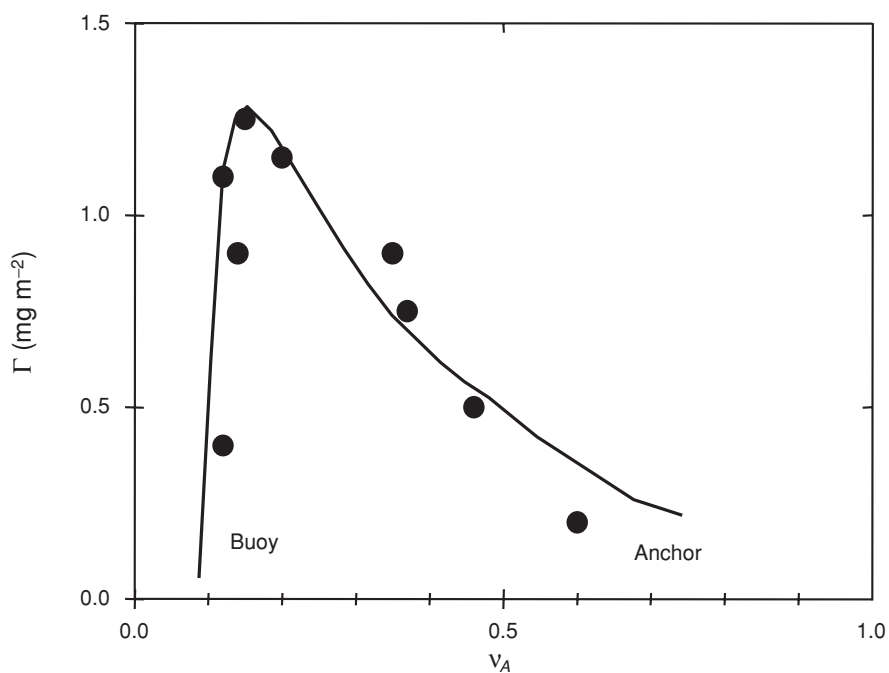
The experimental data in Figure 7.32 are for a series of ABA block copolymers of PEO and poly(propylene oxide) (PPO) on PS latex which show clearly the trend of increasing adsorbed amount up to a certain critical fraction of the adsorbing monomer. The general results are quite typical of block copolymer systems.



**Figure 7.30** Theoretical thickness or adsorbed amount for an adsorbed random copolymer based on the Scheutjens–Fleer theory.

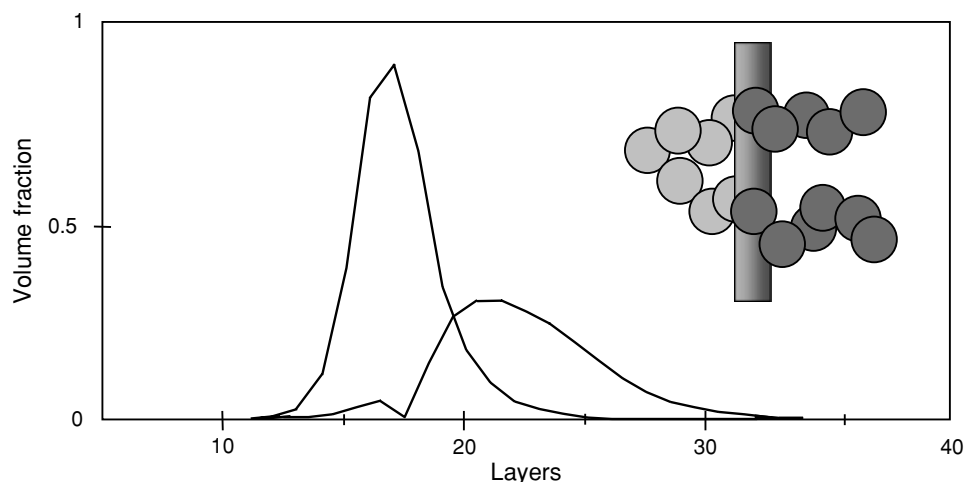


**Figure 7.31** Theoretical thickness or adsorbed amount for an adsorbed block copolymer based on the Scheutjens–Fleer theory.



**Figure 7.32** The adsorbed amount for a series of pluronic (ABA) block copolymers adsorbed on polystyrene latex.

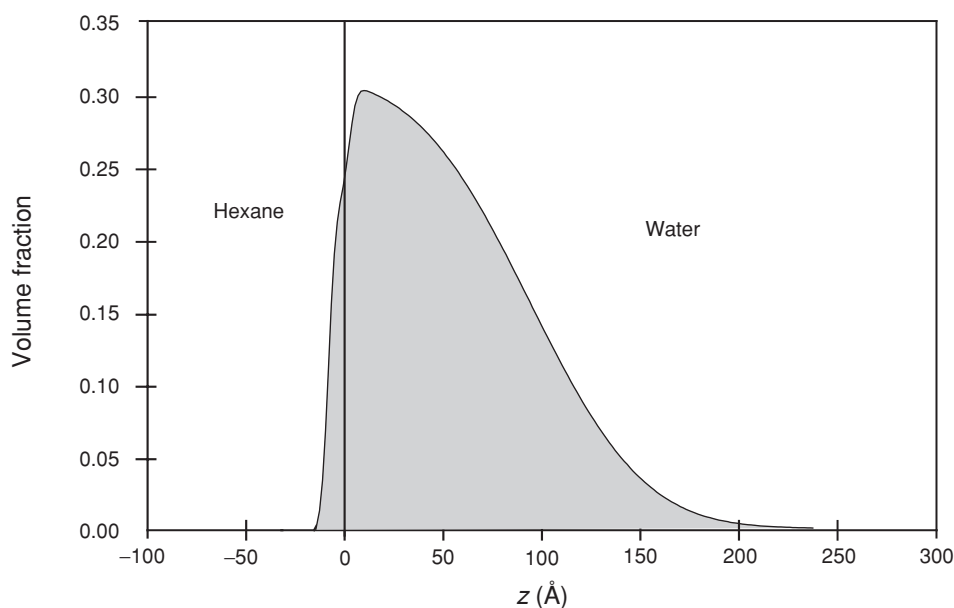




**Figure 7.33** SCF calculation for the adsorption of an ABA block copolymer at the liquid/liquid interface. The inset shows schematically how the segments could arrange.

### 7.5.1 Liquid/liquid interfaces

The same formalism can be used to investigate block copolymers adsorbed at penetrable interfaces, for example an emulsion where the hydrophobic block penetrates into the oil phase and the hydrophilic one into water. Figure 7.33 shows how this partition can take place. The experimental system shown in Figure 7.34 is for a pluronic PEO-PPO-PEO block



**Figure 7.34** Neutron reflectivity derived volume fraction profiles for an ABA block copolymer of PEO and PPO adsorbed at the hexane water interface.

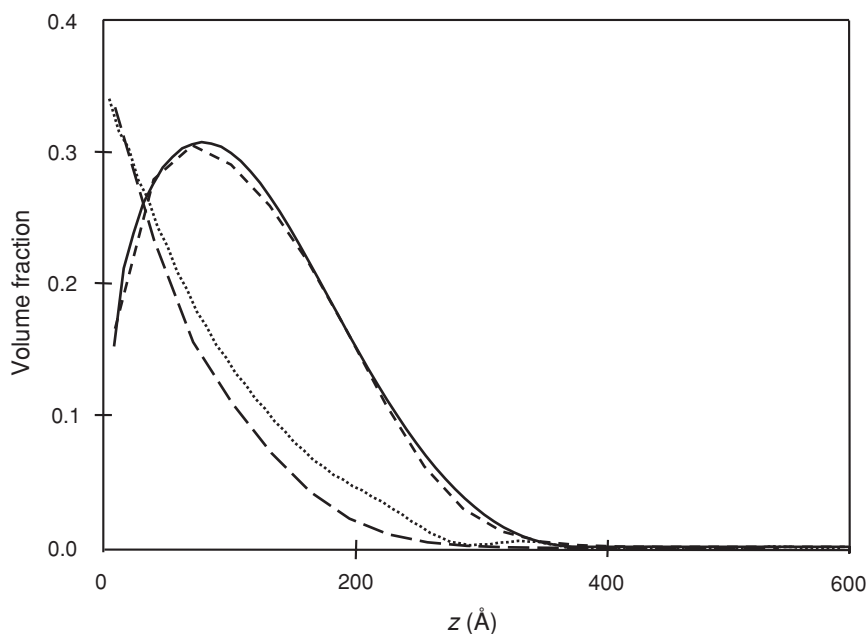
copolymer with 96 monomers in each of the two PEO chains and 69 PPO units. The data were obtained using neutron reflection. In this instance the PPO is not very soluble in the oil and the penetration is not as large as would be found, for example, with an alkyl chain.

## 7.6 Polymer Brushes

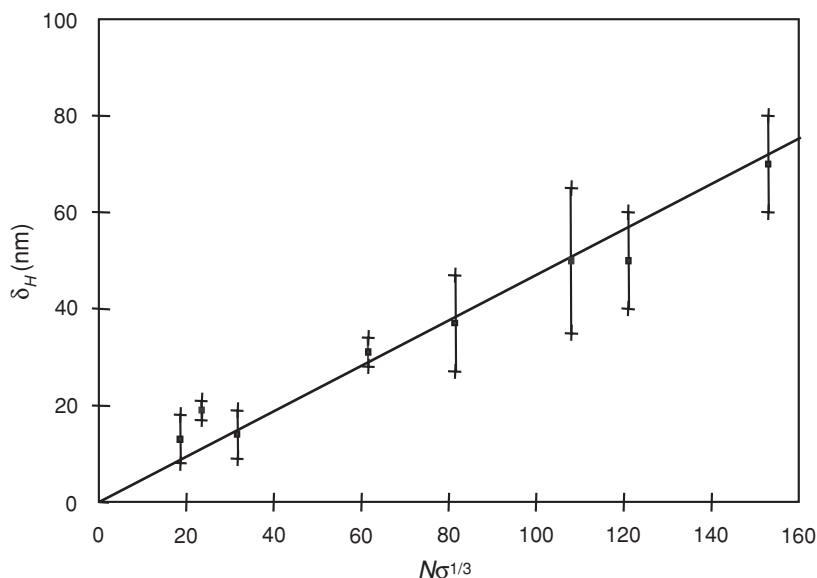
The terminally attached chain discussed at the beginning of this section can also be investigated using the basic SF method and for comparison Figure 7.35 shows examples similar to the ones in Section 7.3.

The brush system has been studied extensively and in Figure 7.35 we show an experimental example of PS grafted onto silica and dispersed in dimethylformamide (DMF) which is a good solvent and from which PS does not physically adsorb. Under these conditions we would expect the mushroom picture. In toluene where physical adsorption does take place, we expect a pancake. The comparison of the SANS data with the SF (solid line and long dashes) prediction is startling.

The same brush system as above has been studied with PCS and SANS to compare with the scaling prediction for a brush, Equation (7.11). From Figure 7.36 it can be seen that for a series of samples of different molecular weights and grafting densities we get



**Figure 7.35** A comparison of the volume fraction profiles obtained from small-angle neutron scattering and SCF calculations for polystyrene chemically grafted to silica. Short dashes and solid lines (theory) correspond to a pure toluene solvent ( $\chi_s = 1.0$ ) and the dotted and long dashes (theory) correspond to the effect of pre-adding dimethyl formamide ( $\chi_s = 0.0$ ).



**Figure 7.36** The hydrodynamic thickness of terminally attached brushes of polystyrene on silica as a function of chain length and grafting density.

a universal plot. The height of the brush surprisingly scales with the number of segments ( $\sim N$ ) which is in marked contrast to a polymer in a good solvent solution where the coil size  $\sim N^{0.6}$ .

## 7.7 Conclusions

In this chapter we have learnt that the structure of a polymer at an interface is very different from that in solution. Two very different scenarios can occur: adsorption and depletion. In the adsorbed state the polymer chain no longer has spherical symmetry and there is an entropic penalty, but this is offset by a gain in energy through adsorption and the gain in entropy on displacing the solvent from the interface. The ends of the adsorbed chain stretch away from the surface by a considerable distance compared to the solution radius of gyration and these farthestmost tail segments are responsible for the particle hydrodynamics and the onset of interactions between particles. Polydispersity can lead to rounded isotherms and effectively irreversible adsorption. Random copolymers can show behaviour intermediate between those of the two respective homopolymers whereas block copolymers can segregate at surfaces giving an anchor and a buoy layer. There is an optimum value of block size to give the maximum thickness for a given overall chain length. For terminally attached chains, mushroom and pancake layer shapes can be found, depending on the attraction of the polymer to the surface, and for this system the thickness is linear in chain length.

For depletion interactions and its effects on colloid stability see Chapter 8.

## References

1. Fleer, G., Stuart, M.C., Scheutjens, J.H.M.H., Cosgrove, T. and Vincent, B. (1993) *Polymers at Interfaces*. Chapman & Hall, London.
2. Jones, R.A.L. and Richards, R.W. (1999) *Polymers at Surfaces and Interfaces*. Cambridge University Press, Cambridge.
3. Netz, R.R. and Andelman, D. (2003) *Phys. Rep.*, **380**, 1–95.
4. Granick, S., Kumar, S.K., Amis, E.J., Antonietti, M., Balazs, A.C., Chakraborty, A.K., Grest, G.S., Hawker, C., Janmey, P., Kramer, E.J., Nuzzo, R.H., Russell, T.P. and Safinya, C.R. (2003) *J. Polym. Sci. B*, **41**, 2755–2793.
5. Adamson, A.W. and Gast, A.P. (1997) *Physical Chemistry of Surfaces*. Academic Press, New York.
6. Binder, K. (1995) *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*. Oxford University Press, Oxford.
7. de Gennes, P.-G. (1979) *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca.
8. Milner, S.T. (1991) *Science*, **251**, 905–914.
9. Marshall, J.C., Cosgrove, T., Leermakers, F., Obey, T.M. and Dreiss, C.A. (2004) *Langmuir*, **20**, 4480–4488.

## Chapter 8

# Effect of Polymers on Colloid Stability

*Jeroen van Duijneveldt*

### 8.1 Introduction

Many colloidal suspensions in practice also contain polymers. These may be adsorbed to the particles, chemically attached to their surfaces or they may be free in solution. Whichever form they take, their presence has a major effect on the stability of colloidal suspensions. This chapter explores some of those possibilities. The effect of polymers on colloid stability is an area of active research interest and a few examples of recent results are given below.

#### 8.1.1 Colloid stability

What is meant by colloid stability? As discussed in Chapter 1, in practice there are several levels at which the stability of suspensions can be considered. First of all, we focus on the interaction between a pair of particles in suspension. Generally van der Waals forces (see Chapter 3) operate between any two colloidal particles in suspension and these can lead to strong attractions at close contact of the particles. If no stabilisation mechanism is provided, the particles will aggregate rapidly (a non-equilibrium process).

Once a stabilisation mechanism has been provided to avoid irreversible aggregation, the suspension as a whole may still have a tendency to phase separate into a dilute and a dense phase for instance, as a result of weak particle attractions. A phase equilibrium may be obtained, or long-lived non-equilibrium (such as gel) states may form.

Yet another meaning of the term colloid stability relates to the tendency of a suspension to undergo sedimentation. If particles are well stabilised in terms of their pair interaction, the stabilised suspension may still undergo sedimentation so the final suspension may not be considered stable. This chapter shows how polymers can be used to affect colloidal stability at all these levels.

#### 8.1.2 Limitations of charge stabilisation

The charge stabilisation of colloidal particles is discussed in detail in Chapter 3. However, there are a number of limitations in relying on charge stabilisation alone. Effective stabilisation requires the use of a polar solvent, for instance. Charge stabilised particles are very sensitive to addition of salt, in particular high valency counter ions. In fact charge stabilised

suspensions are essentially unstable (there is only a kinetic barrier to aggregation), which can create difficulties at high concentrations or when the suspensions are sheared. Finally, if a suspension is likely to undergo freeze-thaw cycles, it may not recover from the freezing cycle.

### 8.1.3 Effect of polymers on interactions

The outline of this chapter is as follows. After an introduction of particle interactions three key mechanisms are discussed by which polymers can affect interactions between colloidal particles, and hence their overall stability:

- Steric stabilisation, where a polymer is adsorbed or chemically attached to the surface of the particles and the presence of the polymer gives rise to particle repulsions
- Depletion interactions, where non-adsorbing polymers in solution induce attractions
- Bridging interactions, where adsorbing polymers induce particle attractions by adsorbing to two particles at the same time.

For more details about colloids and in particular the role of polymers in controlling colloid stability, the following textbooks are recommended. The main text on polymeric stabilisation is still the monograph by Napper [1]. Some more recent developments are summarised in the texts by Fleer and co-workers [2] and by Jones and Richards [3]. The description of steric interactions below is based on the approach presented by Russel *et al.* [4]. Surface forces are discussed in detail in the book by Israelachvili [5].

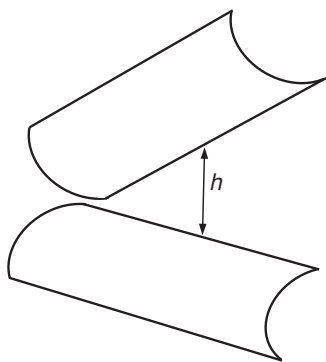
## 8.2 Particle Interaction Potential

First we consider the total interaction potential between two particles. This may consist of a number of different contributions. For instance, van der Waals attractions will in general have to be taken into account. In particular in polar solvents (but not exclusively), charge interactions need to be considered. Finally there are interactions due to the presence of polymer chains. In the general case all these may be of importance simultaneously. However in this chapter we will focus on the polymer mediated interactions alone. By studying suspensions close to refractive index matching, van der Waals attractions are minimised (see Chapter 3). This is most easily realised in non-aqueous systems.

Due to their high molecular weight, polymer chains have a large number of degrees of freedom and structural correlations between polymer segments complicate their theoretical description. This chapter only includes a few simplified approaches to gain a qualitative understanding, whilst providing references to more detailed descriptions. A further point to note is that the timescales involved in the structural relaxation of polymers in solution can be significant, which means that colloidal particles do not necessarily experience the equilibrium interaction potential described by theory. Examples of this are given below.

### 8.2.1 Measuring surface forces

In developing a fundamental understanding of the role of polymers in the interaction between surfaces (and hence particles) a key role was played by the surface force apparatus (SFA), developed by Israelachvili and co-workers (see [5]). In the SFA two crossed cylindrical



**Figure 8.1** Crossed cylinder geometry used in surface force apparatus.

mica surfaces (see Figure 8.1) are used, with mica being chosen because this material can be cleaved to give a molecularly smooth surface.

The surfaces are immersed in a solution and brought very close together. They are mounted on a stiff cantilever and an interferometric technique is used to determine the distance  $h$  between the two surfaces, and the deflection of a laser beam is used to detect the bending of the cantilever, from which the surface force can be obtained.

This technique gives detailed information on the interaction as a function of the distance between two surfaces, or nanometre length scales. It measures  $F(h)/R$ , where  $R$  is the mean radius of curvature of the mica cylinders. The Derjaguin approximation then relates this force to the interaction energy between flat surfaces per unit area,  $E(h)$ , as

$$F(h)/R = 2\pi E(h) \quad (8.1)$$

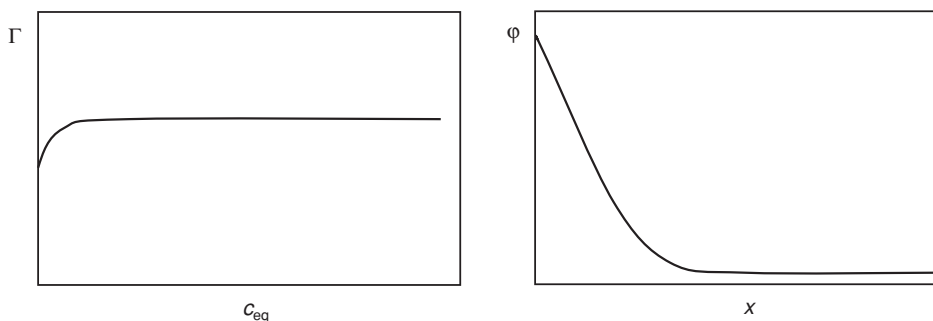
In recent years several other methods have been developed to probe forces between surfaces of which two are mentioned here. In total internal reflection microscopy (TIRM) the interaction between a colloidal particle and a flat surface is studied by recording the particle–surface distance distribution [6]. Atomic force microscopy (AFM) was developed to characterise surfaces but it has also been applied to study interactions between a colloidal particle and a flat surface [7]. One advantage of the latter method is that dynamical interactions can be probed as well, i.e. the role of approach velocity can be examined.

### 8.3 Steric Stabilisation

An important mechanism for providing colloid stability is that of steric stabilisation. It involves covering the colloidal particles with a dense polymer layer. In a good solvent for the stabilising polymer, this gives rise to steep repulsions between the particles.

#### 8.3.1 Theory

Perhaps the most common method of coating particles with a polymer layer is to allow the polymer to adsorb from solution. A detailed description of the resulting adsorbed layer is given in Chapter 7. Polymers at a surface may either adsorb or deplete, depending on the value of the surface  $\chi$  parameter  $\chi_s$ . If it is larger than a critical value,  $\chi_s > \chi_{sc}$ , then

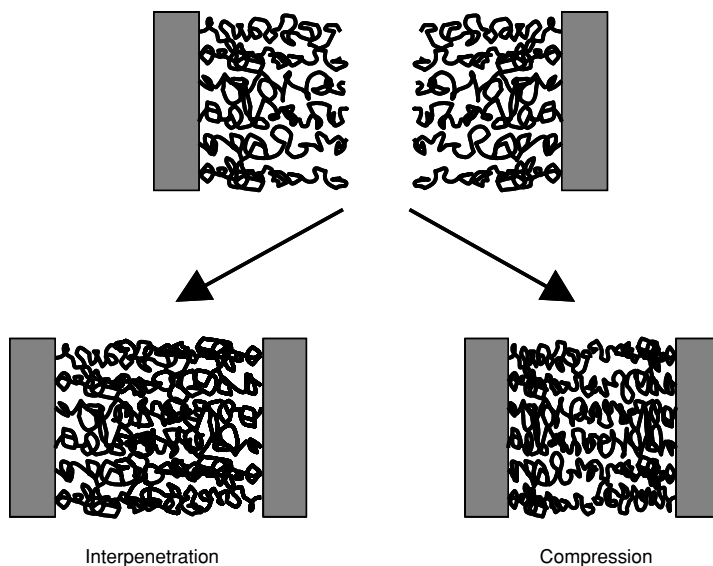


**Figure 8.2** Summary of polymer adsorption behaviour. The left hand panel shows an adsorption isotherm and the right hand panel shows the polymer volume fraction profile.

adsorption will result, and if it is less the polymer will be depleted from the surface instead. Figure 8.2 summarises the adsorption process.

The left hand panel shows an adsorption isotherm, giving the adsorbed amount per unit of surface,  $\Gamma$ , as a function of the equilibrium concentration of polymer in solution,  $c_{eq}$ . As the equilibrium concentration is increased, in this case  $\Gamma$  quickly reaches a plateau. In practice this plateau is typically a few milligrams per square metre of surface of particles. In the right hand panel the polymer segment density is sketched as a function of distance from the surface,  $x$ , summarising the structure of the polymer layer. Predicting such properties analytically is not easy but there are numerical methods available which we can use to calculate such properties, for instance the Scheutjens–Fleer method detailed in Chapter 7.

Figure 8.3 illustrates the principle of steric interactions. The particles are covered with a dense layer of polymer, either through adsorption or through chemical grafting. As a first



**Figure 8.3** Principle of steric interactions.



approximation each polymer layer is represented as having a constant polymer segment density  $n$  throughout the layer.

When two particles approach, the polymer layers will touch as soon as the particle surface distance  $h$  becomes less than twice the layer thickness. Whilst maintaining the polymer concentration in each layer as a step function, two possible scenarios now arise: (a) interpenetration, in which the two polymer layers gradually intermingle as the surfaces approach, locally doubling the polymer segment concentration, and (b) compression, in which the concentration of polymer segments between the surfaces gradually increases from its initial value but is at the same value across the gap between the surfaces.

In a good solvent (Chapter 6) for the polymer the local increase in polymer concentration carries a free energy penalty and this gives rise to a repulsive steric interaction between the particles. This repulsion will be felt as soon as the surface separation  $h$  becomes less than twice the unperturbed layer thickness  $\delta$ ,  $h < 2\delta$ , and it will increase steeply as soon as  $h < \delta$ , when compression of the polymer layers is inevitable. This steep repulsion provides the steric stabilisation of suspensions. Whereas van der Waals attractions still operate between the colloidal particles, effective stability may ensue provided the attractions are weak (compared to  $k_B T$ ) at separations  $h \approx \delta$ , where the steric repulsion sharply increases (see also Chapter 3).

A simple analytical expression for the steric interaction energy  $V_{\text{ster}}$  between two polymer coated particles of radius  $a$  was derived by Fischer [8]:

$$V_{\text{ster}}/k_B T = 4\pi a \Gamma^2 N_A \frac{\bar{v}_2^2}{\bar{V}_1} \left( \frac{1}{2} - \chi \right) \left( 1 - \frac{h}{2\delta} \right)^2 \quad (8.2)$$

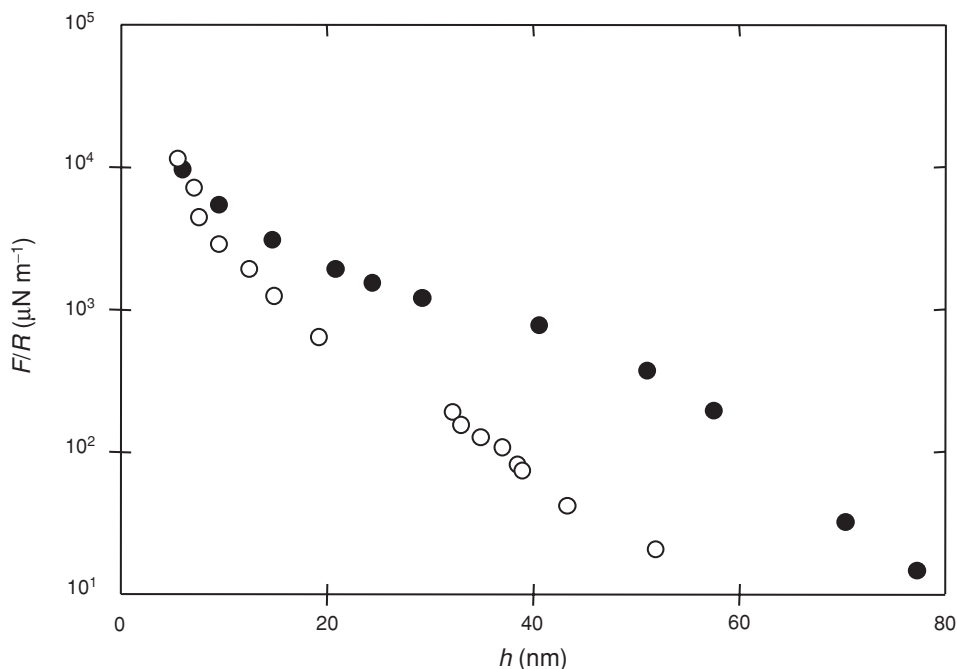
The derivation assumes a constant segment density in the polymer layers and assumes linear superposition of these densities and is therefore only valid for weak overlap between the layers, in the regime  $\delta < h < 2\delta$ . The partial specific volume of the polymer chains is written  $\bar{v}_2$  and the molar volume of the solvent molecules as  $\bar{V}_1$ . In this expression the interactions scale as the square of the surface coverage  $\Gamma$ , thus highlighting the importance of a high surface coverage for effective steric stabilisation. For repulsive interactions,  $V_{\text{ster}} > 0$ , good solvent conditions are required, i.e. the Flory–Huggins parameter  $\chi < 0.5$ .

The role of solvent quality will be explored further in Section 8.6. A detailed overview of theories of steric stabilisation is given in the text by Napper [1].

In the case of adsorbing polymers the interaction will also depend on the equilibrium conditions chosen, i.e. whether the adsorbed amount is able to adjust itself during the approach of the surfaces (full equilibrium) or not. The text by Fleer *et al.* [2] gives a much more detailed account of the different possibilities and of theoretical approaches, such as the Scheutjens–Fleer method, to tackle them.

An example of steric interactions between two surfaces due to adsorbed polymer is shown in Figure 8.4. A surface force apparatus was used to study the force–distance curve between mica sheets to which poly(ethylene oxide) (PEO) was adsorbed with  $M_w = 160,000 \text{ g mol}^{-1}$ , from an aqueous solution containing  $10 \text{ mg L}^{-1}$  PEO and  $0.1 \text{ M KNO}_3$ . For this polymer  $R_g = 13 \text{ nm}$  and adsorption was essentially irreversible (after Klein and Luckham [9]).

Initial compression follows the upper curve (full circles). Repulsions become visible at a surface separation of  $6R_g$ . Rapid decompression followed by compression follows the lower curve (open circles). The force is now less than before, which is ascribed to the slow



**Figure 8.4** Force–distance curve for PEO ( $M_W = 160,000 \text{ g mol}^{-1}$ ) in  $0.1\text{M KNO}_3$ . Full circles: initial compression and open circles: compression after rapid decompression (after [9]).

adjustment of the adsorbed polymer concentration profile between the plates (on a typical time scale of min–hour).

### 8.3.2 Steric stabiliser design

In order to achieve effective steric stabilisation, a number of conditions need to be met. For adsorbing polymers we require:

- High surface coverage,  $\Gamma$
- Strong adsorption,  $\chi_s > \chi_{sc}$
- Good solvent for stabilising chain,  $\chi < 0.5$
- Low free polymer concentration,  $c_{eq} \approx 0$ .

Strong repulsions are only obtained at high surface coverage, and in fact at low surface coverage bridging interactions might otherwise result (see below). The adsorbing polymer needs to be in the plateau region of the adsorption isotherm (Figure 8.2). Strong adsorption is required to ensure high surface coverage, and good solvent condition is needed for effective repulsions. Finally a low concentration of free polymer is desirable to avoid the depletion attractions discussed below. Although it is sometimes possible to achieve successful steric stabilisation using homopolymers, a more promising route in general is to employ copolymers.

In practice copolymers are used, built up from two different monomers, which we denote here as A, which adsorbs strongly to the particles and B, which does not adsorb and dissolves well in the solvent. Different architectures can be considered:

- Random copolymers A/B
- BAB block copolymers
- A-(B)<sub>n</sub> graft copolymers
- Chemically grafted polymer B
- Surfactant with tail group B.

Although an A/B random copolymer is straightforward to synthesise, this architecture is not very effective at creating a thick, dense stabilising layer. More effective is to use a BAB-type block copolymer where the A portion would adsorb and the two B portions would stick out into solution and provide steric stabilisation. Farther along the same line, an A-(B)<sub>n</sub> type graft copolymer can be used, where the A backbone adsorbs and the B portions would stick out into solution again. Finally, B-type polymers can be end-grafted onto the particle surfaces. The advantage is that the polymer is then bound covalently and hence will not desorb. Removal of the unreacted polymer, if required, is time consuming, however. Finally, surfactants may be used in the same way, where the head groups of the surfactants would adsorb to the particles, and the tail groups of the surfactants would provide the steric stabilisation layer.

Model suspensions in which the suspended particles have near-hard sphere interactions can be obtained by relying on steric interactions, using a polymer layer which is thin compared to the particle core. Two well-studied systems are (a) poly(methylmethacrylate) (PMMA) stabilised with a graft copolymer of poly(hydroxystearic acid) (PHS) [10] and (b) silica stabilised with a grafted layer of stearyl alcohol [11]. In both cases non-aqueous solvents are used, with near-refractive index matching of the particles with the solvent minimising Van der Waals attractions.

### 8.3.3 Marginal solvents

As mentioned above, for effective steric stabilisation a good solvent for the stabilising polymer chains is required. In other words, steric stabilisation is expected to break down around the  $\Theta$  temperature (where  $\chi = 0.5$ ) [1]. An example of practical importance is the stability of PEO stabilised particles in water, where aggregation may be observed on temperature increase (depending on salt concentration) as aqueous solutions cease to be a good solvent for PEO at elevated temperatures.

To describe the interactions between polymer layers the approach of Russel *et al.* [4] is followed here. It starts from a model due to de Gennes and Alexander, representing each polymer layer as having a constant polymer concentration, i.e. a step function of distance (more realistic density profiles are presented in Chapter 7). In this approach the Helmholtz free energy of a polymer solution of a polymer of  $N$  statistical segments of length  $l$  is written as

$$\frac{A}{Mk_B T} = \ln \frac{n}{N} - 1 + \frac{1}{2} N v n + \frac{1}{6} N w n^2 \quad (8.3)$$

where  $n$  is the polymer segment density,  $M = nV/N$  is the number of polymer chains per unit volume,  $w^{1/2}$  is the physical volume per polymer segment and the so-called excluded

volume per segment  $v = w^{1/2}(1-2\chi)$  takes into account the solvent quality. Starting from this an approximate expression is derived for the free energy per chain in a grafted polymer layer of thickness  $\delta$ ,

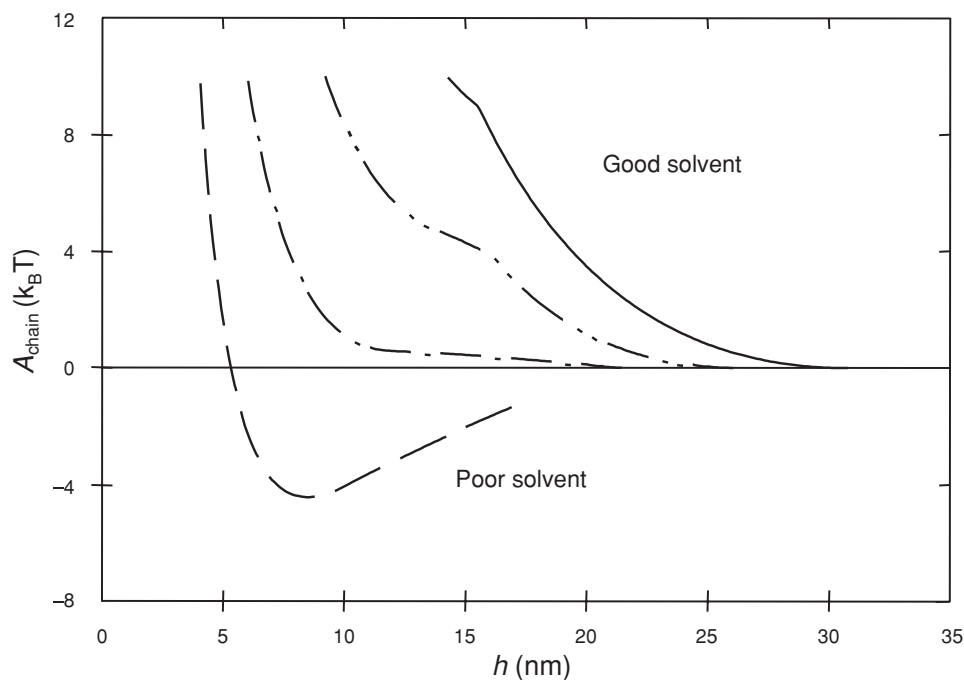
$$\frac{A_{\text{chain}}}{k_B T} = \frac{3}{2} \left( \frac{\delta^2}{Nl^2} + \frac{Nl^2}{\delta^2} - 2 \right) + \frac{1}{2} Nvn + \frac{1}{6} Nwn^2 \quad (8.4)$$

The thickness of a polymer layer can be determined from this equation under the constraint that the total amount of polymer in the layer is fixed,  $n\delta = NTN_A/M_w$ .

In the approach by Russel *et al.* linear superposition of two polymer layers is again assumed which allows the steric interaction energy per chain to be evaluated as a function of the surface separation  $h$ . To illustrate the role of solvent quality in steric interactions this theory is used here to obtain a few specific results. Figure 8.5 shows predicted interaction curves (expressed as interaction energy per polymer chain,  $A_{\text{chain}}$ ).

The calculations were carried out to model the behaviour of a suspension of silica spheres (diameter 88 nm) to which polystyrene (PS) of  $M_w = 26,600 \text{ g mol}^{-1}$  was end-grafted to a moderately high coverage,  $\Gamma = 2.3 \text{ mg m}^{-2}$ . The particles were suspended in cyclohexane, a marginal solvent for PS [12]. In this solvent the  $R_g$  of PS of this size is 4.4 nm.

Figure 8.5 shows interaction curves for different values of the  $\chi$  parameter. In a very good solvent ( $\chi = 0$ ) steeply repulsive interactions are predicted which already set in at a surface



**Figure 8.5** Effect of solvent quality on steric interactions obtained from mean field theory (polystyrene stabilised silica, see text for details; solvent  $\chi$  parameter (top to bottom) 0.0, 0.31, 0.57, 0.79).

separation of 30 nm. On approaching  $\Theta$  solvent conditions the stabilising layers shrink somewhat and as a result repulsions only set in at smaller separation  $h$ , but still rise steeply at short separations. Finally, in this case for  $\chi \approx 0.6$ , at intermediate distances an attractive interaction is obtained. The fact that purely repulsive interactions are still predicted under slightly worse than  $\Theta$  conditions is due to the high physical packing fraction of polymer segments in the stabiliser layers (represented by  $w$  in the theory) compensating for the negative excluded volume (represented by  $v$ ).

Experimentally the particles discussed here were found to aggregate at temperatures a little below the  $\Theta$  temperature for PS in cyclohexane of 34°C [12]. At low particle concentrations fractal, open aggregates were formed, whereas space-filling gels were obtained at particle volume fractions around 0.1.

Theoretical curves such as those in Figure 8.5 suggest that one would have a rather delicate control of particle interactions, through the temperature. Once they have aggregated, it should be possible to redisperse particles by reheating, and thus improving solvent quality again. Indeed with short stabilising chains (stearyl alcohol grafted silica for instance) this tends to be the case. However particles treated with longer stabiliser chains (such as the PS treated particles discussed here) are often hard to redisperse after they have been aggregated for some time. This is possibly a result of the polymer chains having formed bridges across to the next particle (see Section 8.5), or a result of physical entanglement of the polymer chains. In Figure 8.4 time dependent effects in steric interactions were also encountered for adsorbed polymers.

## 8.4 Depletion Interactions

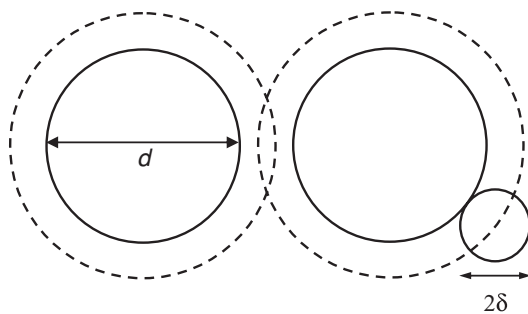
Addition of free (non-adsorbing) polymer in solution induces so-called depletion interactions between colloidal particles. To gain an understanding of this we will discuss the Asakura–Oosawa (AO) model of depletion interactions. The particles are considered as hard spheres of diameter  $d$  and the polymers are represented by little spheres of diameter  $2\delta$ . Within this description the polymer coils do not interact, and hence the osmotic pressure of the polymer solution,  $\Pi$ , can be calculated from their number concentration  $n_{\text{pol}}$  using the van't Hoff law

$$\Pi \approx n_{\text{pol}} k_B T \quad (8.5)$$

The polymer coils do have hard sphere interactions with the colloidal particles, however, and hence they are excluded from a depletion layer with thickness  $\delta$  around each particle, shown as dashed circles in Figure 8.6.

When two particles approach to a surface separation less than  $2\delta$ , the depletion layers overlap and the available free volume for the polymer is increased. Due to the polymer osmotic pressure this results in an effective attraction between the particles ('attraction through repulsion'). The size of the polymer molecules therefore sets the range of these attractions. Their strength can be controlled by varying the polymer concentration. Following this argument the depletion potential is given by

$$V_{\text{dep}} = -\Pi V_{\text{ov}}, \quad (d < r < d + 2\delta) \quad (8.6)$$



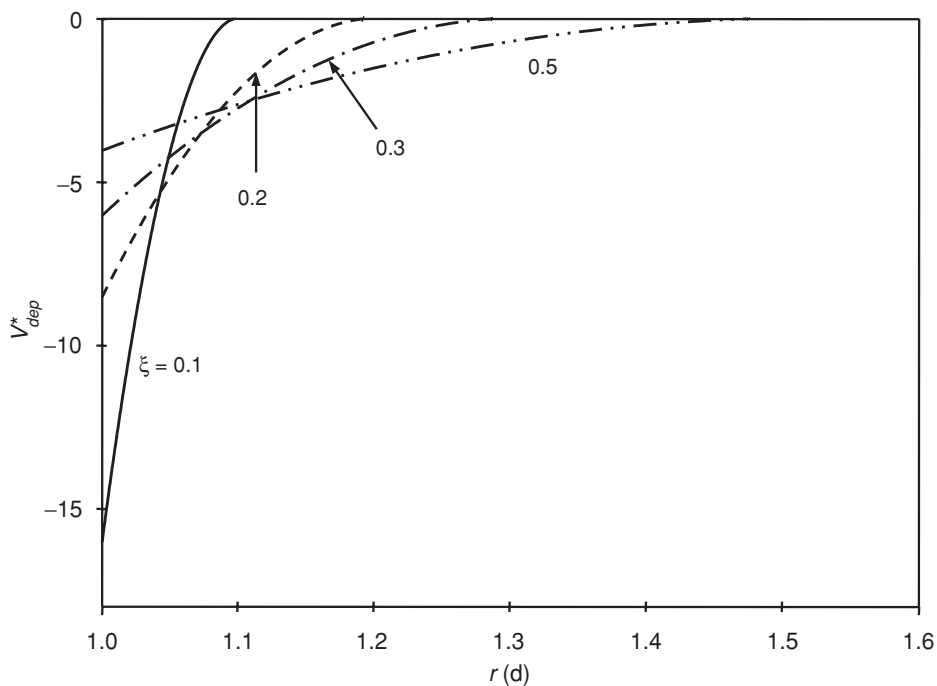
**Figure 8.6** Asakura–Oosawa model for depletion interactions.

where  $r$  is the particle centre–centre distance and the overlap volume of the depletion layers,  $V_{ov}$ , is given by

$$V_{ov} = \left( 1 - \frac{3r}{2d(1+\xi)} + \frac{1}{2} \left[ \frac{r}{d(1+\xi)} \right]^3 \right) \frac{\pi}{6} d^3 (1+\xi)^3 \quad (8.7)$$

with  $\xi = \delta/a$  being the polymer/colloid size ratio.

Examples of this potential are shown in Figure 8.7 as  $V_{dep}^* = V_{dep}/\Pi v_0$  for a few values of  $\xi$ . The interaction potential is normalised by dividing by the osmotic pressure and by



**Figure 8.7** Depletion potential (Equation (8.6)) for different polymer/colloid size ratios  $\xi$ .

the volume swept out by a polymer coil,  $v_0 = 4\pi\delta^3/3$ , and the particle separation is normalised by the particle diameter,  $d$ . An almost triangular attraction function is obtained, which becomes narrower as the polymer size is reduced. Typically for the size of the polymer coil (thickness of the depletion layer) the radius of gyration of the polymer is taken ( $\delta = R_g$ ) [4].

Although the AO model is rather approximate it has led to useful predictions, some of which are discussed below. A more detailed description of depletion interactions is given in [2]. This is an area of active research, see for instance [13] and references therein.

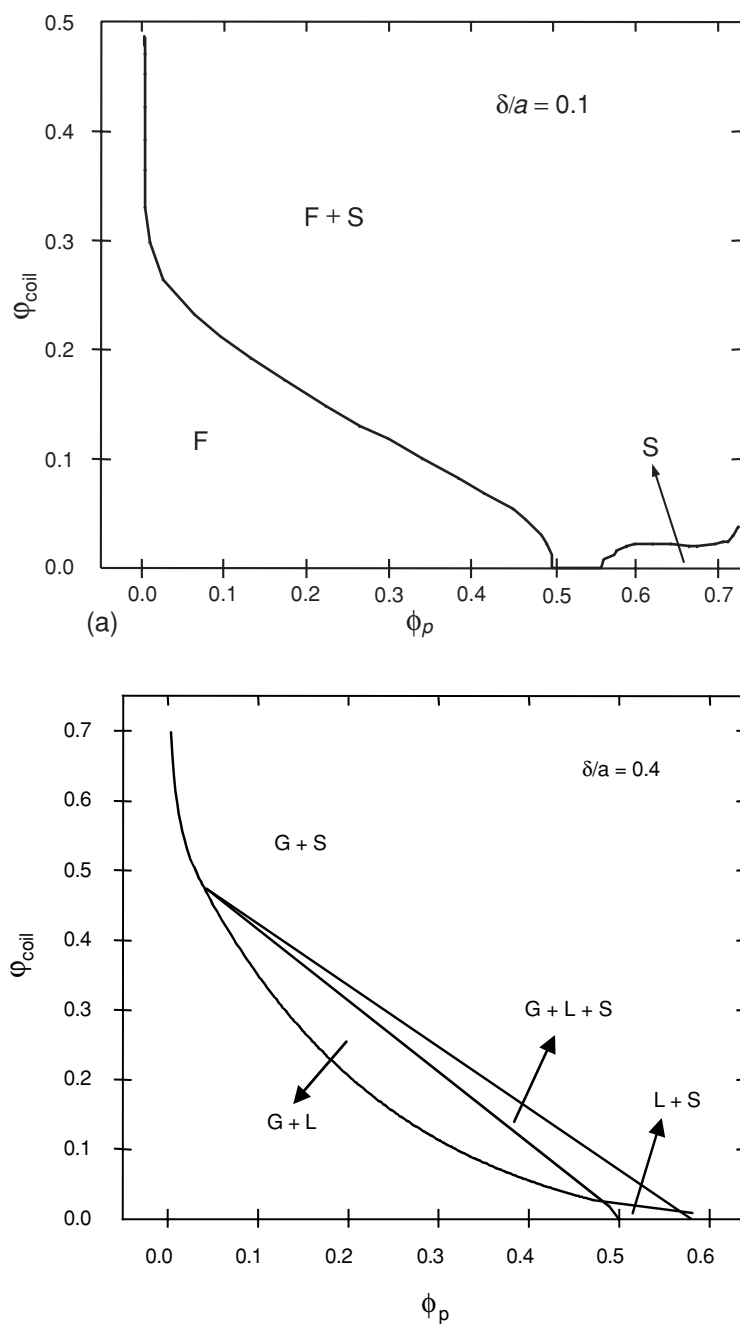
The depletion mechanism allows for the switching on of particle attractions in a very controlled manner and the range and strength of the attractions can be varied independently. As a result phase transitions can take place where different phases are formed, which differ in the concentration and ordering of the colloidal particles. A few theoretical predictions are shown here, obtained by combining the AO model with scaled particle theory for the polymer free volume [14]. In Figure 8.8 calculated phase diagrams are plotted for size ratios  $\xi = 0.1$  and  $\xi = 0.4$ . The polymer concentration is expressed as the volume fraction of coils,  $\phi_{\text{coil}} = v_0 n_{\text{pol}}$ . Monodisperse hard spheres form colloidal crystals at particle volume fractions  $\phi_p > 0.5$  so the pure colloidal suspension already has a fluid to solid transition. This was verified experimentally using sterically stabilised PMMA colloids [15]. It is assumed here that the particles have a narrow size distribution; if the relative polydispersity exceeds about 10% crystals are not obtained.

On addition of a small polymer (compared to the colloid) (short range attractions,  $\xi = 0.1$ ) a wide immiscibility gap opens up, with a dilute fluid coexisting with a dense solid. Using somewhat larger polymers (long range attractions,  $\xi = 0.4$ ) a qualitatively different behaviour emerges with a gas – liquid transition at intermediate polymer and colloid concentrations, and an area where all three phases (gas, liquid and solid) coexist. Note that the requirement of equal chemical potentials implies that not only the colloid but also the polymer concentration is different in the three phases.

Experimental work confirms these predictions [16]. However, experimentally at high attraction strength (polymer concentration) non-equilibrium states are often found, such as aggregates (at low particle concentration) or gels (at higher particle concentration). These non-equilibrium states can be long lived and may actually be the state a product is formulated in (for instance a weak gel). At moderate polymer concentration, the resulting gels may be weak enough to undergo syneresis, with a sudden collapse observed in some cases [17].

An important point to note about depletion interactions is that low polymer concentrations, often only a few  $\text{mg ml}^{-1}$ , can induce phase separation. When sterically stabilised particles are used with a thick stabilising polymer layer, the depletion effects are significantly reduced as the free polymer is now able to penetrate the stabilising layer. Also, the depletion effect may be lost on increasing polymer concentration further, resulting in restabilisation [2].

The AO model discussed here is suitable to describe what is referred to as the ‘colloid limit’ with  $\xi < 1$ . However one can also consider the opposite limit,  $\xi > 1$  where small particles are added to solutions of large polymers. This is relevant for instance in protein crystallisation experiments where it is common practice to add polymer in solution to aid the crystallisation. Hence this is referred to as the ‘protein limit’. Phase separation in dilute (gas) and dense (liquid) phases can still occur. However, a different theoretical description is required as the polymer concentrations involved are typically above the overlap concentration,  $\phi_{\text{coil}} > 1$  [18, 19].



**Figure 8.8** Theoretical phase diagrams for colloid–polymer mixtures based on the AO model. Fluid (F), solid (S), gas (G) and liquid (L) phases are predicted to form. Size ratio (a)  $\xi = 0.1$  and (b)  $\xi = 0.4$  (after [14]).



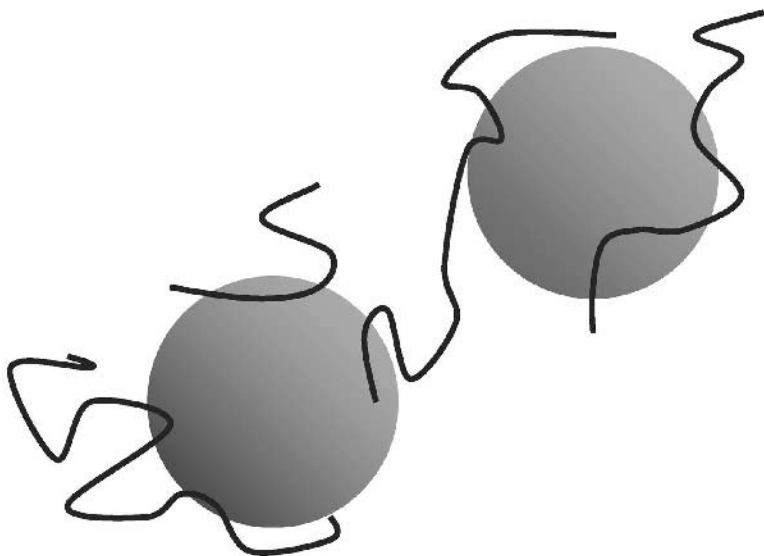
This chapter focuses on the role of polymers in controlling interactions between colloidal particles. However, as an aside it is worth noting that depletion type interactions occur more generally. In suspensions containing more than one suspended species it can be useful to think of one (typically the smallest) species as inducing a depletion attraction between the larger particles. The depletant can for instance consist of small particles, or perhaps of surfactant micelles. However, a theoretical description of such systems is more difficult than the idealised polymer case discussed here, because interactions between the small objects themselves usually cannot be neglected.

An example is that of hard sphere mixtures, which for size ratio larger than 5 are predicted to demix as a result of such depletion interactions. For such a phase separation to occur high overall concentrations are typically needed, in the region of 50% by volume of solids [20].

## 8.5 Bridging Interactions

Finally we consider the case of bridging interactions (see Figure 8.9). At low surface coverage  $\Gamma$ , adsorbing polymers may attach themselves to the surface of more than one particle. The effect of this on the particle interactions is referred to as the bridging interaction. Whereas steric interactions are repulsive under good solvent conditions, the bridging interaction can lead to attractions also under better than  $\Theta$  conditions when homopolymers are adsorbed [4, 5].

Polymeric flocculants can be designed to maximise bridging interactions in cases where aggregation is desired. ABA block copolymers where the A block adsorbs are types of molecules which lend themselves for this type of behaviour. Often polyelectrolytes are used, where charges on the polymers are used to enhance adsorption to the target particles. Counter



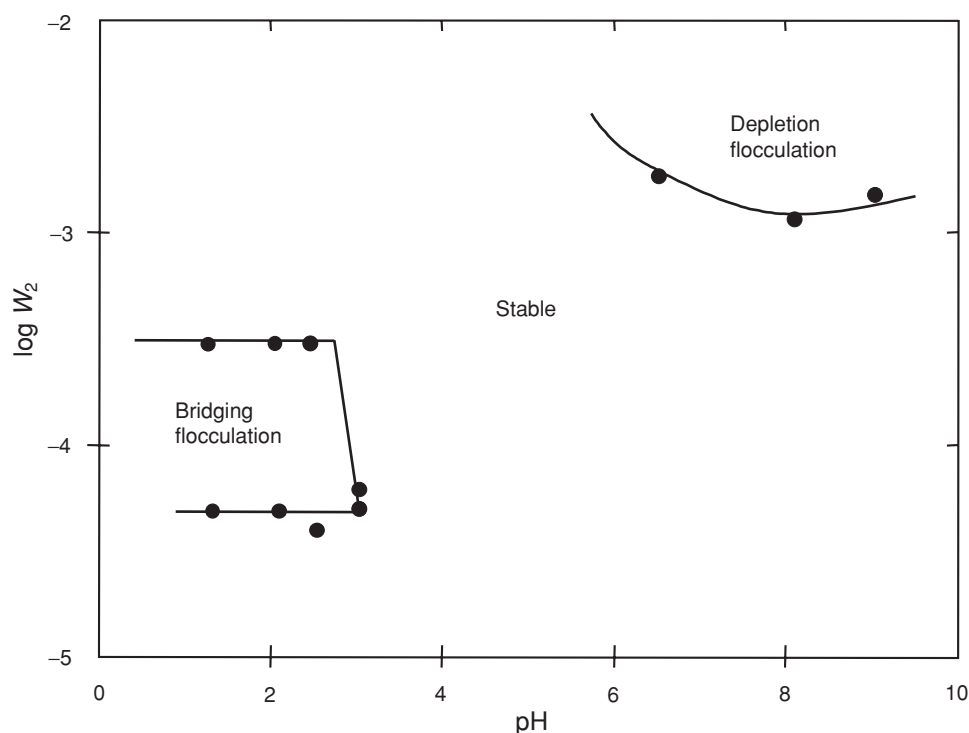
**Figure 8.9** Principle of bridging interactions.

intuitively, it is possible to achieve bridging flocculation also with polymers with the same charge sign as the particles; the adsorption is then mediated by counter ions such as  $\text{Ca}^{2+}$ .

Polymeric flocculants are used in a variety of applications, for instance waste water treatment, clarification of drinks such as beer and wine and also in processing of minerals. Following aggregation, suspended fine particles will settle down more easily or the resulting flocs may be filtered off.

So far the effects of steric interactions, depletion and bridging have been discussed separately here. In practice more than one mechanism may have to be taken into account. The same adsorbing polymer may give rise to both steric and bridging interactions, or a weakly adsorbing polymer may cause both bridging and depletion attractions for instance.

As an illustration of the complex scenarios possible, see Figure 8.10. A stability map is shown for aqueous polystyrene (PS) latex particles, carrying terminally grafted poly(ethylene oxide) (PEO) chains, to which poly(acrylic acid) (PAA) is added ( $M_w = 14,000 \text{ g mol}^{-1}$ ), as a function of pH [21].  $W_2$  is the weight fraction of added PAA. At low pH the PAA coacervates with the PEO and at low PAA concentration this results in bridging flocculation. At higher pH the PAA does not adsorb onto the PEO coated particles, but instead depletion flocculation may occur at sufficiently high concentrations of added PAA. An intermediate range of pH values is observed where neither effect is sufficient to cause flocculation.



**Figure 8.10** Stability map of PS latex particles stabilised with PEO chains on addition of PAA as a function of pH.  $W_2$  is the weight fraction of PAA (after [21]).

## 8.6 Conclusion

Many products and processes involve colloidal suspensions and often polymers are also involved in solution or attached to the particles. The presence of these polymers has a major impact on the aggregation state and therefore on the flow properties (rheology, see Chapter 11) of these suspensions. Whereas the influence of polymers on the behaviour of solid, spherical particles is discussed in this chapter, the types of interaction are general and would also apply to the behaviour of emulsion droplets or of non-spherical (rod-like or plate-like) particles.

Depending on the application, there is a range of possibilities for what the practical aim might be. When a stable suspension of finely suspended small particles is required, as in inks, a good steric stabilisation is needed. However, larger particles would typically have a tendency to settle down even if well-stabilised in terms of their particle interaction potential. Equivalently oil-in-water emulsion droplets would tend to cream up. A weak flocculation induced by depletion interactions might help prevent this settling or creaming by building a weak gel network. Finally there are applications such as water clarification where a rapid and strong aggregation of the particles is needed, and polymeric flocculants could be used.

## References

1. Napper, D.H. (1983) *Polymeric Stabilization of Colloidal Dispersions*. Academic Press, New York.
2. Fler, G.J., Stuart, M.A.C., Scheutjens, J.M.H.M., Cosgrove, T. and Vincent, B. (1993) *Polymers at Interfaces*. Chapman and Hall, London.
3. Jones, R.A.L. and Richards, R.W. (1999) *Polymers at Surfaces and Interfaces*. Cambridge University Press, Cambridge.
4. Russel, W.B., Saville, D.A. and Schowalter, W.R. (1989) *Colloidal Dispersions*. Cambridge University Press, Cambridge.
5. Israelachvili, J. (1992) *Intermolecular & Surface Forces*. 2nd ed. Academic Press, London.
6. Prieve, D.C. and Frej, N.A. (1990) Total internal-reflection microscopy – a quantitative tool for the measurement of colloidal forces. *Langmuir* 6(2), 396–403.
7. Milling, A. and Biggs, S. (1995) Direct measurement of the depletion force using an atomic-force microscope. *J. Colloid Interface Sci.*, 170(2), 604–606.
8. Fischer, E.W. (1958) *Kolloid Z.*, 160, 120.
9. Klein, J. and Luckham, P.F. (1984) Forces between two adsorbed poly(ethylene oxide) layers in a good aqueous solvent in the range 0–150 nm. *Macromolecules* 17, 1041–1048.
10. Antl, L., Goodwin, J.W. Hill, R.D., Ottewill, R.H., Owens, S.M. and Papworth, S., *et al.* (1986) The preparation of poly(methyl methacrylate) lattices in nonaqueous media. *Colloids Surf.*, 17(1), 67–78.
11. van Helden, A.K., Jansen, J.W., and Vrij, A. (1981) Preparation and characterization of spherical monodisperse silica dispersions in non-aqueous solvents. *J. Colloid Interface Sci.*, 81(2), 354–368.
12. Weeks, J.R., van Duijneveldt, J.S. and Vincent, B. (2000) Formation and collapse of gels of sterically stabilized colloidal particles. *J. Phys. Condens. Matter*, 12(46), 9599–9606.
13. Tuinier, R., Rieger, J. and de Kruif, C.G. (2003) Depletion-induced phase separation in colloid-polymer mixtures. *Adv. Colloid Interface Sci.*, 103(1), 1–31.
14. Lekkerkerker, H.N.W., Poon, W.C.K., Pusey, P.N., Stroobants, A. and Warren, P.B. (1992) Phase-behavior of colloid plus polymer mixtures. *Europhys. Lett.*, 20(6), 559–564.

15. Pusey, P.N. and van Megen, W. (1986) Phase-behavior of concentrated suspensions of nearly hard colloidal spheres. *Nature*, **320**(6060), 340–342.
16. Ilett, S.M., Orrock, A., Poon, W.C.K. and Pusey, P.N. (1995) Phase-behavior of a model colloid-polymer mixture. *Phys. Rev. E*, **51**(2), 1344–1352.
17. Starrs, L., Poon, W.C.K., Hibberd, D.J. and Robins, M.M. (2002) Collapse of transient gels in colloid-polymer mixtures. *J. Phys. condens. Matter*, **14**(10), 2485–2505.
18. Bolhuis, P.G., Meijer, E.J. and Louis, A.A. (2003) Colloid-polymer mixtures in the protein limit. *Phys. Rev. Lett.*, **90**(6), 068304.
19. Sear, R.P. (2001) Phase separation in mixtures of colloids and long ideal polymer coils. *Phys. Rev. Lett.* **86**(20), 4696–4699.
20. Dijkstra, M., van Roij, R. and Evans, R. (1999) Phase diagram of highly asymmetric binary hard-sphere mixtures. *Phys. Rev. E*, **59**(5), 5744–5771.
21. Cawdery, N., Milling, A. and Vincent, B. (1994) Instabilities in dispersions of hairy particles on adding solvent-miscible polymers. *Colloids Surfaces A*, **86**, 239–249.

# Chapter 9

## Wetting of Surfaces

*Paul Reynolds*

### 9.1 Introduction

The sciences of surfaces and colloids are inextricably linked since the chemistry and physics governing the surface properties of both large and small areas are identical in that they describe the interactions between molecules and collections of molecules. As a result the interaction between a material placed on a surface and an extensive solid surface can be adequately described using our knowledge of both surfaces and colloids. In this chapter we outline the established theories and explanations of wetting and use these to provide a more practical appreciation of aspects of these phenomena. The basic principle is simple in that when a liquid is placed on a solid surface it will spread to some extent. Here both liquid and solid can be described as bulk. However, in the region of interaction where the bulk phases meet, local interactions are of primary importance. Consider the adsorption of a macromolecule or a surfactant or indeed a grease or oil stain on a surface and the effect that this has on a liquid placed on the surface. Clearly this surface modification, made with only a molecular or macromolecular dimension, has a dramatic influence on the macroscopic observation of the liquid drop behaviour on that surface. Wetting, the coverage of a surface with another material, is an important industrial and academically interesting and often challenging area.

The basic premise developed in this chapter is that a liquid in contact with a solid in the presence of a given vapour exhibits behaviour which depends on all three components. The rules that govern this behaviour may not be well developed and are empirical in many instances. However, there is a scientific rationale behind these observations. It is the intention here to illustrate how these rules are developed and expose the scientific background to the interactions.

### 9.2 Surfaces and Definitions

We would normally associate wetting with the coverage of a solid surface with a liquid material. In doing so we would assume that the liquid is mobile on the solid surface. The liquid surface is a dynamic surface whereas the solid surface is not. Interestingly, perhaps what can be thought of as a fundamental property of a solid and a liquid, the surface energy and the surface tension, respectively, have identical units i.e. they are dimensionally the same. An

interesting discussion of the units and dimensions is given in Adamson and Gast [1]. It is important to note that vapour has not been omitted from the definitions of surface tension and surface energy and is a very important component of the overall interaction. Moreover, given the necessity to include vapour it becomes clear that this will change depending upon the temperature, pressure and partial pressure of the vapour phase.

There are four surfaces that we can consider in the absence of a vacuum:

- Liquid–vapour    surface tension
- Liquid–liquid    interfacial tension
- Liquid–solid    surface energy or interfacial tension
- Solid–vapour    surface energy

The units of both surface tension and surface energy are Newton per metre ( $\text{N m}^{-1}$ ) and Joule per square metre ( $\text{J m}^{-2}$ ). Although dimensionally the same and equivalent in value, they are defined differently as stated below.

Each of the surfaces identified above contributes to the overall picture of the wetting of a surface, which in the case of a liquid on a solid leads to a well defined shape of droplet for any given materials and conditions. In describing this behaviour we shall determine which properties of the materials are important in controlling wetting.

### 9.3 Surface Tension

Consider first the liquid–vapour surface. It is a very dynamic surface where molecules from the liquid phase are leaving the surface and molecules from the gaseous phase are entering the liquid phase continuously. The interface is a difficult region to model and requires the development of models based upon an imaginary surface known as the Gibbs dividing surface. This gives the basis for developing equations for surface excesses (see Section 4.3.2). Unfortunately many elementary texts prefer to show a liquid schematically as a set of spheres with interactions radiating in all directions and being equivalent to one another. At the surface of the liquid we lose the equivalence in the direction of the surface, simply because there are no molecules above the surface. As a result it is ‘shown’ that the force, or tension, arising in a surface is due to the ‘missing forces’ of interaction of a molecule in the surface of the liquid phase in the direction of the vapour phase. The picture is schematically satisfying but lacking in scientific rigor, yet surface tension can be shown to depend only on the interactive forces between molecules.

However, we can make measurements in the surface which relate to surface properties. These lead to definitions of the surface tension. It is possible to write two definitions of surface tension and both will take some mental agility to fully appreciate what is meant. Thus common definitions are: *surface tension is the force acting at right angles to a line of unit length in the surface of a liquid*, or *surface tension is the force per unit length, which when multiplied by the distance moved to create a new area is equal and opposite to the work done in extending the area of surface by unity*.

The latter expression shows that the surface tension can be interpreted as the energy per unit surface area and that this has a tendency to reduce its area to a lower free energy arrangement. A full discussion of surface tension is given in Chapter 4.

## 9.4 Surface Energy

We can think about a surface as being created by the work required in bringing a molecule from the bulk to the surface. Thus, *the work that is required to increase that surface area by unity is the surface free energy*. Of course there is a tendency of that surface to contract and this is where the picture of molecules in the surface at least illustrates the ideas here because we can visualise that state of tension in the surface and we would define that as the surface tension.

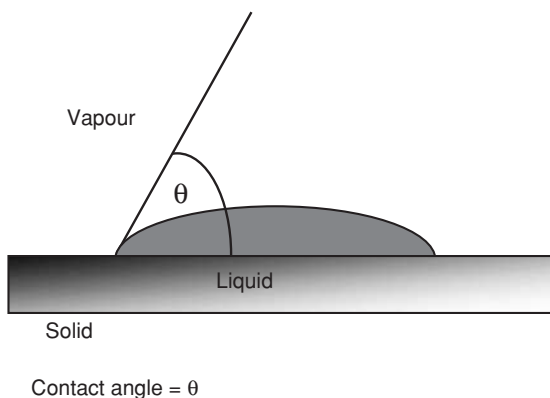
## 9.5 Contact Angles

When a liquid is placed on a solid surface it can spread to form a continuous film or form discrete droplets. In the case of the latter a range of different behaviour can be observed. It has been stated previously that for a given solid and liquid and a defined set of conditions, e.g. temperature and pressure, a liquid drop will form a well defined shape on the solid surface. Figure 9.1 shows this schematically but additionally shows that the shape subtended by the liquid at the three-phase line of contact, where the solid (s), liquid (l) and vapour (v) meet, has a defining angle, called the contact angle ( $\theta$ ).

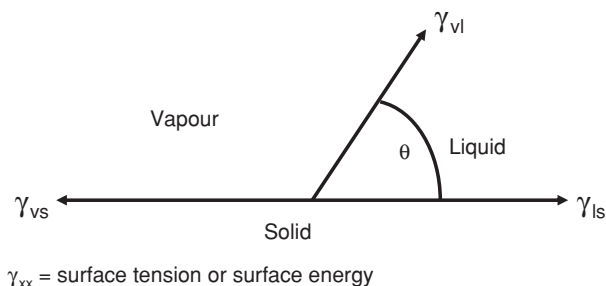
Interpretation of this into more fundamental forces reveals the nature of the contact angle. This is done by considering the tensions in each of the surfaces as shown in Figure 9.2.

Taking the tensions at the three-phase point of contact and putting them into a balance of forces resolved in the plane of the solid surface, it is seen that there are three tensions ( $\gamma$ ): one acting in the direction of the vapour/solid ( $\gamma_{vs}$ ) and this is opposed by the two tensions, the surface tension in the liquid–solid surface ( $\gamma_{ls}$ ) and a component of the surface tension of the vapour/liquid ( $\gamma_{lv}$ ). When resolved in the plane of the solid surface the function becomes the cosine of the contact angle. The resolution of these tensions gives Young's equation [2]

$$\gamma_{ls} + \gamma_{lv} \cos \theta = \gamma_{vs} \quad (9.1)$$



**Figure 9.1** Contact angle at the three-phase line of contact of solid, liquid and vapour.



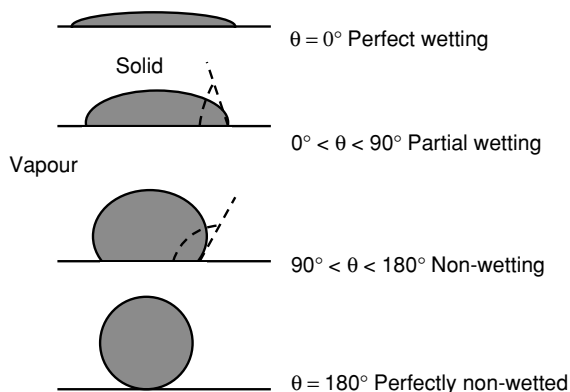
**Figure 9.2** Equilibrium balance of forces for the tensions of liquid drop on a solid surface.

It must be remembered however, that the contact angle, whilst giving useful and fundamental information about the solid/liquid interaction, is defined in the presence of the vapour phase. This is often forgotten and a change in vapour (partial) pressure and composition can have very profound effects on the subsequent contact angle.

## 9.6 Wetting

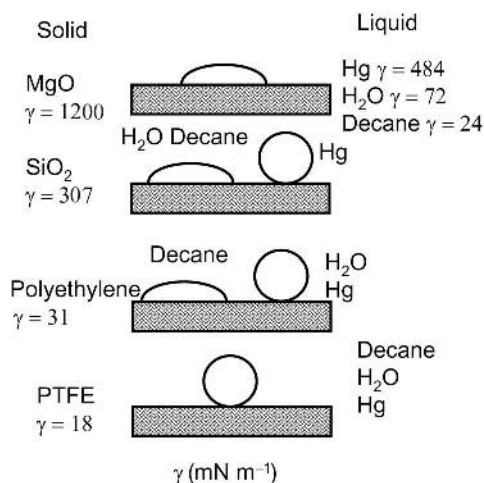
Armed with the previous definitions, wetting *per se* can now be discussed. For some liquids a zero contact angle is obtained and might also be called perfect wetting and hence spontaneous spreading. Another possibility is partial wetting, where a contact angle is subtended somewhere between  $0^\circ$  and  $90^\circ$ . The  $90^\circ$  may well be thought of as an arbitrary distinction between wetting and non-wetting but nonetheless we find it an important distinction. An angle subtended between  $90^\circ$  and  $180^\circ$  in the liquid would be a non-wetting condition and finally of course if the contact angle is  $180^\circ$  then we have a perfectly non-wetted surface. These cases are shown schematically in Figure 9.3.

When discussing wetting behaviour it is most often convenient to speak in terms of contact angle. The general principle of wetting can be defined by the characteristic surface properties of the solid and liquid as shown in Figure 9.4.



**Figure 9.3** Wetting as described by different values of the contact angle.





**Figure 9.4** Differential wetting of three liquids on solid surfaces of different surface energy.

Consider three different liquids, mercury, water and decane, having surface tensions of 484, 72 and 24  $\text{mN m}^{-1}$  respectively, that are placed on a planar surface of a material. In the first instance we choose a high energy surface, magnesium oxide with a surface energy of 1200  $\text{mN m}^{-1}$ . We would expect a finite contact angle and therefore wetting. However, the three liquids would show a different interaction with a solid with a reduced surface energy. For example if we use silica with a surface energy of 307  $\text{mN m}^{-1}$ , water and decane would be seen to wet the surface whereas mercury would not. Thus mercury, with a higher surface tension than the silica surface energy, does not wet the surface whereas water and decane do. Reducing the energetics of the solid surface further, using polyethylene with a surface energy of 31  $\text{mN m}^{-1}$ , shows that only decane wets the surface; mercury and water do not. Again we will see that mercury and water have higher surface tensions than the surface energy of polyethylene. Finally, a low energy solid surface, polytetrafluoroethylene (PTFE), shows non-wetting of all three liquids, and all three liquids, have higher surface tensions than the surface energy of PTFE alone.

It is tempting, using the above observations, to conclude that a useful rule is: *liquids with low surface tensions wet solids with high surface energies.*

Evidently, the magnitude of the contact angle follows some rationale governed by the magnitude of surface tensions and energies. A more complete description of the governing scientific principles is beyond the scope of this chapter; however, it would entail a consideration of Hamaker constants which describe the interactions between molecules in adjacent materials [3] (see Section 4.3.2). This idea is adopted later empirically to describe the observed behaviour.

There are other generalisations which start to appear when studying comparative contact angles measured for liquids in contact with materials. For example, Table 9.1 identifies some materials in contact with a variety of liquids.

From the above table it can be seen that water has a range of contact angles when in contact with a low energy solid surface like PTFE going through to higher energy solid surfaces like glass. There is a clear trend in these values.

**Table 9.1** Liquid surface tension and contact angles on solid surfaces

Liquid	$\gamma$ (mN m <sup>-1</sup> )	Solid	$\theta$ (degree)
Hg	484	PTFE	150
Water	73	PTFE	112
		Paraffin wax	110
		Polyethylene	103
		Human skin	75–90
		Gold	66
		Glass	0
Methyl iodide	67	PTFE	85
		Paraffin wax	61
		polyethylene	46
		PTFE	46
Benzene	28	Graphite	0
		PTFE	40
<i>n</i> -decane	23	PTFE	30
<i>n</i> -octane	2.6	PTFE	170
Tetradecane/water	50.2	PTFE	

There are other trends apparent in these data and one can be seen for PTFE; high contact angles (150°) are found for mercury, through to *n*-octane which has a contact angle of 30°. The difference in the surface tensions of the liquids ranges from 484 mN m<sup>-1</sup> for mercury to 2.6 mN m<sup>-1</sup> for *n*-octane. The above observations can be used to conclude another tentative rule: *high surface tension liquids tend to have large contact angles, while polar solids tend to have smaller contact angles.*

## 9.7 Liquid Spreading and Spreading Coefficients

The idea that there is a rational organisation to our observations can be taken a little further with the specific case of perfect wetting (zero contact angle).

It is apparent from Young's equation (9.1) that when the contact angle,  $\theta$ , is zero:

$$\gamma_{sv} - \gamma_{ls} - \gamma_{lv} = 0 \quad (9.2)$$

For real wetting we can evaluate (9.2) as the spontaneous spreading coefficient,  $S$ . If  $S$  is positive then spontaneous wetting occurs whereas a negative value of  $S$  would result in a finite contact angle.

The coefficient  $S$  has some interesting properties when examined practically. One example is the wetting properties of the homologous series of paraffins in contact with a given material. As the carbon number is changed  $S$  can become progressively less negative until spreading occurs. For example hexane spreads on water but decane does not. Perfect wetting

occurs at the point at which the surface energy and surface tension are equivalent and this observation can be used for surface characterisation [4].

## 9.8 Cohesion and Adhesion

Surface energies can also be used to define cohesion and adhesion from the condition of spontaneous spreading. Adhesion and cohesion simply define the interactions which lead to the cohesive nature of materials and the adhesive nature of one material when in contact with another material. There is no implication in these definitions of any chemical bonding. Adhesion between a solid and liquid is defined as the work required to separate the solid from the liquid. Thus, the work of adhesion involves the creation of new surfaces of vapour/liquid and solid/vapour, and the destruction of the old surface of liquid/solid. The Dupré equation defines this and Figure 9.5 shows a schematic of the process:

$$W_{ls} = \gamma_{sv} + \gamma_{vl} - \gamma_{ls} = W_a \quad (9.3)$$

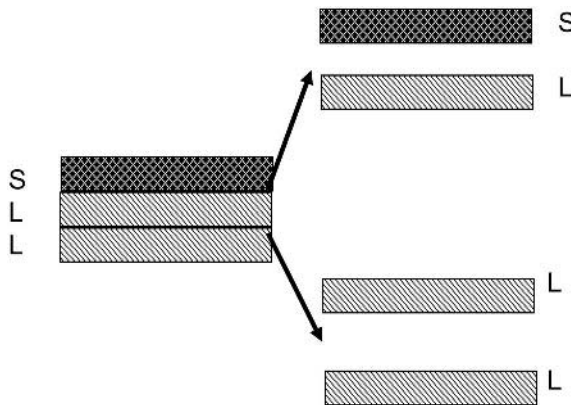
where  $W_a$  is the work of adhesion and  $W_s$  is the work required to separate the solid from the liquid surface. It is worth noting that all the quantities are defined per unit surface area.

In the same way cohesion can be seen to be the creation of two liquid/vapour interfaces and the destruction of a liquid/liquid interface. The cohesive nature of a liquid (or solid) involves the separation of the liquid from itself. Clearly from this construction the work of cohesion,  $W_c$ , can be seen to be

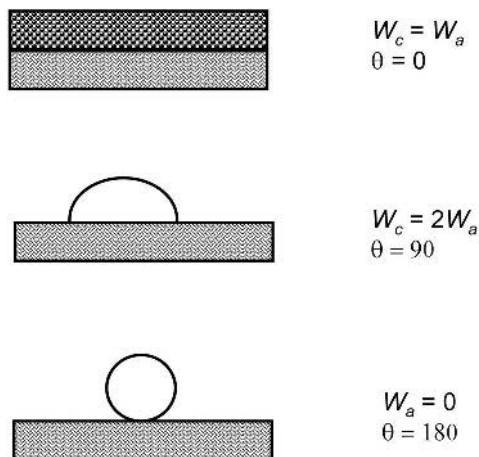
$$W_c = 2\gamma_{lv} \quad (9.4)$$

Adhesion and cohesion have another quite simple relationship which can be developed by taking into account the contact angle, shown as

$$\cos \theta = -1 + 2 \left( \frac{W_a}{W_c} \right) \quad (9.5)$$



**Figure 9.5** Solid-liquid adhesion and cohesion. Adhesion is the separation of the solid from the liquid. Cohesion is the separation of the liquid from itself.



**Figure 9.6** Simple relationships between the work of cohesion and the work of adhesion for a solid and a liquid.

The result of this is that the contact angle is seen to be governed by the competition between cohesion of a liquid to itself and adhesion of a liquid to a solid. Schematically Figure 9.6 shows some simple results. At a  $0^\circ$  contact angle the work of cohesion equals the work of adhesion whereas at a  $90^\circ$  contact angle, the condition just between wetting and non-wetting, the work of cohesion equals twice the work of adhesion. For perfect non-wetting the work of adhesion is zero.

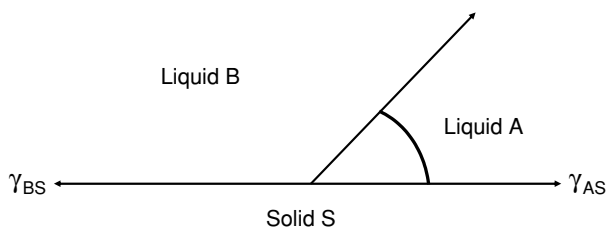
Cohesive failures and adhesive failures in materials are more easily appreciated with this understanding.

## 9.9 Two Liquids on a Surface

So far we have considered the wetting of a surface by a single liquid in equilibrium with its vapour. There is, however, a commercially important case where two immiscible liquids sit on a solid surface and differentially wet the surface. This case is shown schematically in Figure 9.7 where the solid S is differentially wetted by liquid A and liquid B.

This is a commercially important case since very often there are two immiscible liquids in contact on a surface and it is not difficult to find examples.

The appropriate tensions and the directions in which they are acting have been drawn in the diagram, from which it is clear that they can be resolved in the same way as before to



**Figure 9.7** Two immiscible liquids on a solid surface.

**Table 9.2** Differential wetting of a surface by two immiscible liquids

Solid	Liquid A	Liquid B	$\theta_{SA}/(\text{degree})$
Al <sub>2</sub> O <sub>3</sub>	Water	Benzene	22
PTFE	Water	Decane	180
Hg	Water	Benzene	100
Glass	Hg	Gallium	0

recover Young's equation. However, this time the equation (Equation (9.6)) applies to the two liquids in contact with each other and in contact with the solid surface:

$$\gamma_{BS} = \gamma_{AS} + \gamma_{AB} \cos \theta_A \quad (9.6)$$

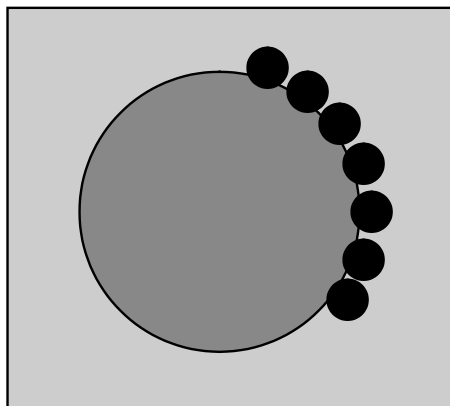
From the diagram it can be seen that liquid A preferentially wets the solid, and that is because the contact angle is smaller for liquid A than it is for liquid B, i.e.  $\gamma_{BS} > \gamma_{AS}$ .

This leads us to another rule of wetting: *for two immiscible liquids on a solid surface, in general, the liquid with the smaller solid/liquid surface tension or surface energy, wets the solid preferentially.*

Table 9.2 shows various results of studying two immiscible liquids wetting the surface of a solid. The solids exemplified here are alumina, PTFE, mercury (note that mercury is also taken as a liquid in the same table) and glass. It can be seen how they are differentially wetted. In the table the angle  $\theta_{SA}$  is the angle subtended by liquid A on the solid surface. Thus in the first case of water and benzene on alumina, liquid A is water having a contact angle of  $22^\circ$  when benzene is present as the second liquid phase. In this case water is better wetting, or the one that is preferentially wetting the solid (alumina). Another example is that of a PTFE surface which can be seen to be completely dewetted by liquid A, water, and thus it will not wet PTFE in the presence of decane. With water and benzene on the surface of mercury, water has a contact angle of  $100^\circ$  and so is dewetted preferentially. These two results may not be surprising in that they show water to be the dewetted phase but it is the magnitude that illustrates that the phenomena are not entirely predicted by our knowledge of a single liquid on a solid surface. The value of  $100^\circ$  for water on mercury, in the presence of benzene, is not far off the  $90^\circ$  distinction between wetting and non-wetting. Complete dewetting,  $180^\circ$ , for water on PTFE in the presence of decane is very rarely observed for a single liquid. Perhaps the most surprising result is that mercury and gallium as liquids in contact with glass show mercury has a zero contact angle and therefore spreads ideally.

A practical example is detergency where initially oil or grease adheres to a fabric surface or perhaps a ceramic surface. It is clear that we need to change the energetics, surface energies, to be able to remove the oil from the surface, and it is generally a water solution that will be used, in conjunction with a surfactant. This case will be considered later.

There are other important commercial cases which can be identified and some occur in the manufacturing industry sector. The preparation of emulsions using small solid particles to stabilise the emulsion necessarily has differential wetting of the solid surface by the two immiscible liquids. These are called Pickering emulsions. Essentially packing of the small particles around the liquid droplet determines the type of emulsion, water in oil (W/O) or oil in water (O/W). The particles therefore stabilise against aggregation and coalescence. The rule is *the liquid phase which better wets the particle is going to be the external phase.*



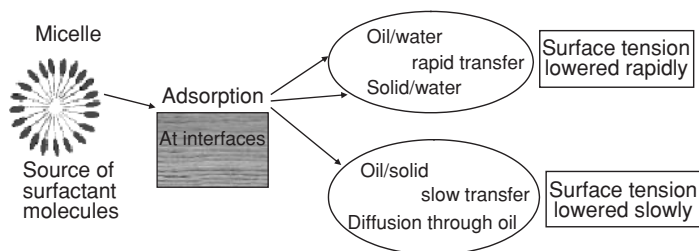
**Figure 9.8** A Pickering emulsion is a liquid droplet stabilised by particles which are better wetted by the continuous phase than the internal phase. More efficient packing leads to a more stable emulsion arrangement.

This we can compare with Bancroft's rule [1], which states that the continuous phase should be the phase in which the surfactant is the most soluble. This is consistent with a simple geometrical argument based on the ratio of the head group area to that of the hydrophobic chain cross-section of a simple surfactant and dictates the type (or stability) of the emulsion (see Figure 5.4). Thus if particles are better wetted by water than oil they will have a greater surface area covered by water. As a result the oriented wedge of water-wetted surface packs more efficiently between a set of spheres surrounding a larger sphere, where the oil only wets a small area of the particles' surface. The result is an oil in water emulsion. This is shown schematically in Figure 9.8. Silica is a good example of this effect whereas carbon black, which is wetted better by oil than water, results in a water in oil emulsion.

There are many other important examples that can be found in the electronics and household products industries. In the process of soldering there is a liquid metal on a surface in contact with a liquid flux, which is there to solubilise impurities and remove them from the solid/liquid metal interface. Clearly this is complex since the temperatures are high and fluctuate and this will lead to changes in the surface energies and the surface tension of the flux. Also the properties of the flux/molten metal interface and flux/metal surface will change as the level of impurities solubilised in the flux changes. In household product applications excess foam production can be a serious problem, for example in washing machines. Antifoamers, which are often liquid silicones containing particulates (e.g. silica), can be used to reduce this [5, 6]. They work by differential surface wetting of the silica by water and silicone which makes the thin foam films unstable.

## 9.10 Detergency

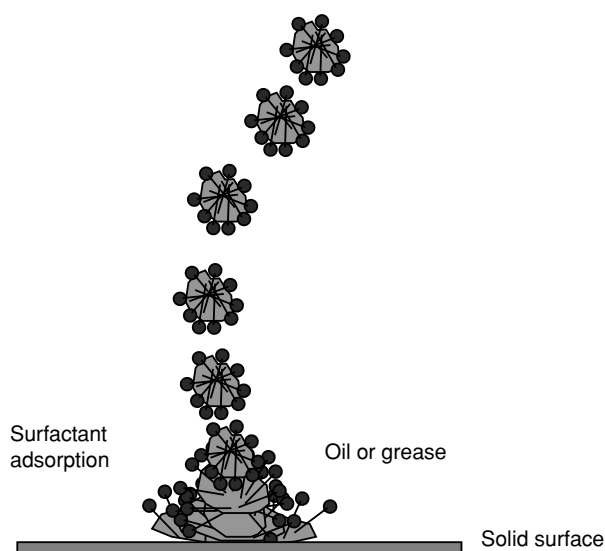
One clear example of wetting of a surface with two immiscible liquids is that of everyday washing. An oil stain (liquid) on a fabric (solid) surface is in contact with an immiscible phase when immersed in washing water. The detergent action at the oil/water/solid interface



**Figure 9.9** Schematic of the detergency process.

is influenced by adsorption of surfactants. The line of tensions may be analysed in much the same way as previously described. The great difference here is that the surface energies and the surface tensions are constantly being modified with time as surfactant from the washing water adsorbs to the oil and fabric surfaces. The dynamics of the mechanism become very important. What is critical, however, is the rate at which the fabric/oil surface is modified by surfactant adsorption. This is a slower process than adsorption at the other surfaces because of the amount of time necessary to allow surfactant to diffuse through the oil droplet. This process leads to the roll-up mechanism leading to the oil droplet detachment. This is shown schematically in Figure 9.9.

The source of surfactant molecules is the micelles in the formulated detergent. These diffuse to the appropriate surfaces and adsorb. The oil drop as it rolls up, because of the contact angle change driven by the change in the surface tension balance of forces, exposes a new surface which then adsorbs surfactant to a greater extent. This stabilises the oil, from re-establishing an adsorbed state on the fabric surface and allows the drop to ultimately detach (Figure 9.10). This process is, however, aided considerably by the agitation occurring



**Figure 9.10** The roll-up mechanism in detergency.

in the washing process, as well as possibly the elevated temperature, although in some parts of the world washing takes place at ambient temperatures and no hot water is used.

## 9.11 Spreading of a Liquid on a Liquid

Spreading can also be observed for two immiscible liquids when one is placed on the surface of the other. The lens formed in this case can be described by a 'generalised Young's equation' (9.7). This is necessary since all the surfaces involved are deformable as shown in Figure 9.11. Under these conditions we need to allow for the contact angles of each to be resolved in a horizontal direction,

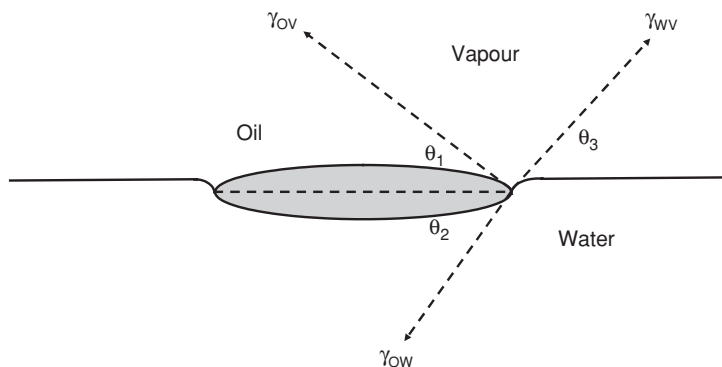
$$\gamma_{wv} \cos \theta_3 = \gamma_{ov} \cos \theta_1 + \gamma_{ow} \cos \theta_2 \quad (9.7)$$

The picture is relatively complicated as spreading develops with time and all the angles change because of the finite size of the spreading liquid reservoir. Again, simply using the definitions and expressions used previously, the balance of forces, leading to work of adhesion ( $W_{AB}$  between liquids A and B) and work of cohesion ( $= 2\gamma_A$  of the spreading liquid), can be used to evaluate the spreading coefficient of one liquid on the other:

$$S = W_{AB} - 2\gamma_A = \gamma_B - \gamma_A - \gamma_{AB} \quad (9.8)$$

Clearly for the spreading of liquid A on liquid B the work of adhesion between A and B must exceed the work of cohesion of the spreading liquid A. The difference between the two quantities is the spreading coefficient,  $S$ , of A on B. As shown previously, if the value of  $S$  is positive spreading results, and if it is negative there is no spreading.

Table 9.3 shows some examples of the spreading of one liquid on another. By taking liquid A as the spreading liquid (octyl alcohol, oleic acid, bromoform, liquid paraffin) on water, we can look at the interfacial tensions and the surface tensions. The surface tension of water in each case is  $72.5 \text{ mN m}^{-1}$ , and the liquid surface tensions of A vary only from 27.5 to 31.8



**Figure 9.11** Surface tensions for a liquid lens on an immiscible liquid.



**Table 9.3** Spreading coefficients of one liquid on another

Liquid A	Liquid A on water ( $\gamma = 72.5 \text{ mN m}^{-1}$ )		
	$\gamma_A \text{ (mN m}^{-1}\text{)}$	$\gamma_{AB} \text{ (mN m}^{-1}\text{)}$	$S \text{ (mN m}^{-1}\text{)}$
Octyl alcohol	27.5	8.5	36.5
Oleic acid	32.5	15.5	24.5
Bromoform	41.5	40.8	-9.8
Liquid paraffin	31.8	57.2	-16.8

$\text{mN m}^{-1}$ . However, the interfacial tensions change dramatically (from 8.5 to  $57.2 \text{ mN m}^{-1}$ ). The spreading coefficients are obtained from the previous relationships and two are positive values (spreading) and two are negative (non-spreading).

Whilst there is a systematic change in spreading coefficients, there are clearly some complicating issues. This is shown by the time-developed behaviour of the spreading of benzene on water. When a drop of benzene is placed on water, a lens forms which expands and spreads. After a period of time it retracts back to reform the original lens. It turns out that there is a limited solubility of benzene in water, and of water in benzene. It takes time for this equilibrium to be established after the drop of benzene is placed on the water surface. From Table 9.4 the initial values of the surface tensions of water and benzene give rise to a spreading coefficient of  $8.9 \text{ mN m}^{-1}$ , the interfacial tension being  $35 \text{ mN m}^{-1}$  and constant. However, after a period of time, water uptake in benzene reduces the surface tension of benzene very slightly, which also increases the spreading coefficient. After a longer period of time the benzene re-equilibrates in the water and the interfacial surface tension of water drops to  $62.2 \text{ mN m}^{-1}$ . The final sum for spreading coefficient is now negative ( $-1.6 \text{ mN m}^{-1}$ ) so it is now non-spreading. The time-dependent dissolution of benzene in water and vice versa alters the behaviour from benzene being spreading to non-spreading on water [1].

The vapour phase composition and (partial) pressures clearly have a significant influence on the spreading behaviour. The role of solubilised impurities in the case of detergency and for two immiscible liquids on a solid surface can also influence the balance of forces markedly. The results may not be as expected. Consider these processes when put into a real situation, the soldering example given above for instance. The fine balance of behaviour to create the desired adhesion/cohesion conditions becomes very difficult to predict.

**Table 9.4** Spreading of benzene on water

	$\gamma_{\text{water}} \text{ (mN m}^{-1}\text{)}$	$\gamma_{\text{benzene}} \text{ (mN m}^{-1}\text{)}$	$S \text{ (mN m}^{-1}\text{)}$
Initial	72.8	28.9	8.9
	72.8	28.8	9
Final	62.2	28.8	-1.6
$\gamma_{WB} = 35 \text{ mN m}^{-1}$			

## 9.12 Characterisation of a Solid Surface

One method of characterising the solid surface is to use a homologous series of liquids of varying surface tension, and observe those which wet and those which do not. Our general rule is that for wetting, the surface energy of the solid/vapour interface is going to be greater than the surface tension of the liquid/vapour interface. For the fluids a ranking order can be arranged which will give a good indication of the solid surface energy from the null spreading condition. However, we can go further than this using the homologous series of liquids, as drops, on the surface.

A plot of the cosine of the contact angle against the liquid surface tension gives us a characterisation tool for the surface. An example is shown in Figure 9.12 and is linear. This is called a Zisman plot [4]. The extrapolation of the cosine of the contact angle for different homologous series tends to the same intercept. This suggests that the surface energy can be characterised by this extrapolated value, which can be termed a 'critical surface tension' ( $\gamma_c$ ). The results, however, should be treated as semi-empirical.

## 9.13 Polar and Dispersive Components

Surface tensions are related to intermolecular forces and an understanding of Hamaker constants shows that they equally relate to intermolecular forces. This is true when dispersion forces dominate, that is to say, when the materials are non-polar. The surface tension properties of liquids can be calculated using Hamaker constants [1, 3] and some values are included in Table 9.5, calculated using the following equation [1]:

$$\gamma_i^d = \frac{A_{ii}}{24\pi (r_i^0)^2} \quad (9.9)$$

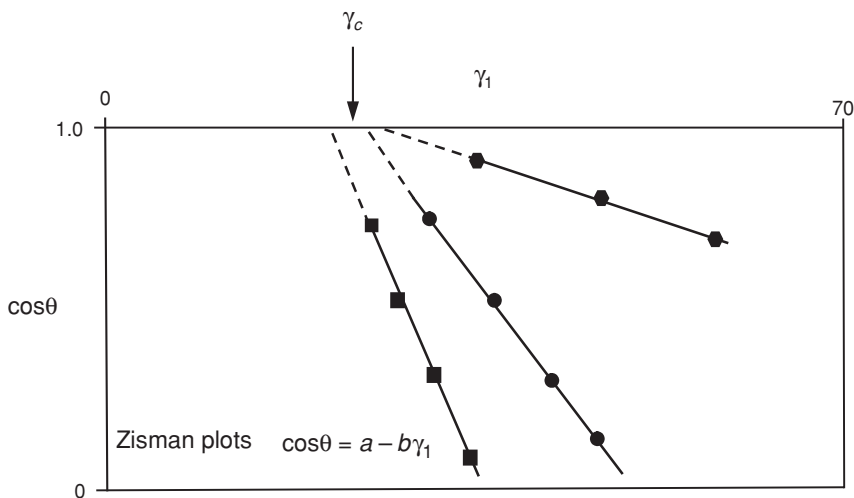


Figure 9.12 Zisman plot for homologous series of liquids on a solid surface.

**Table 9.5** Comparison of calculated (calc) and experimentally (exp) measured values of surface tension

Material	$\gamma_{\text{calc}}^{\text{d}}$ (mN m <sup>-1</sup> )	$\gamma_{\text{exp}}^{\text{d}}$ (mN m <sup>-1</sup> )
<i>n</i> -octane	21.9	21.6
<i>n</i> -hexadecane	25.3	27.5
PTFE	18.5	18.3
Polystyrene	32.1	33
Water	18	72.4

$A_{ii}$  is the Hamaker coefficient and  $r_i^0$  is a distance parameter used in integrating the Hamaker expression [7].

Based only on the dispersion force, it is clear that for non-polar materials (*n*-octane, *n*-hexadecane, PTFE and polystyrene) the predicted values obtained from calculations match almost exactly the experimentally obtained values. When the material has a very polar nature, such as water however, there is a huge disparity between the calculated value and the experimental values.

Furthermore, we can also calculate contact angles as in Equation (9.13) [7]:

$$\cos\theta = -1 + 2\phi \left( \frac{\gamma_{\text{s}}}{\gamma_{\text{lv}}} \right)^{1/2} - \frac{\pi_{\text{sl}}}{\gamma_{\text{lv}}} \quad (9.10)$$

$\phi$  is a constant which accounts for the relative molecular sizes, polarity, etc. and depends upon the molecule under consideration, but under conditions of non-polarity approximates to 1. There is an additional term in this expression and that is surface pressure  $\pi_{\text{sl}}$ , but essentially reduces the solid surface energy through adsorption of vapour from the liquid at the solid/vapour interface. Some caution must be exercised using these equations since significant polar forces clearly affect the results.

## 9.14 Polar Materials

Polar materials can be treated in a similar manner by adding extra contributions to the total surface tension [8]. The surface tensions can be divided into separate independent terms, the most simple being the addition of the polar component and the dispersive component,  $\gamma^{\text{p}}$  and  $\gamma^{\text{d}}$ ; thus

$$\gamma = \gamma^{\text{p}} + \gamma^{\text{d}} \quad (9.11)$$

It is also possible, if necessary, to break the surface tension down into other components, hydrogen bonding for instance. We could add the hydrogen bonding term to the dispersive term, and add in a polar term to that and so on. The total surface tension is subdivided into individual contributions. For a non-polar material, of course, the surface tension is simply equal to the surface tension of the dispersion component.

An alternative approach is to use a set of semi-empirical expressions. These generally contain a term in, or derived from, the geometric mean theorem which is used to combine the Hamaker constants of different materials. Equation (9.12) is a useful expression which

**Table 9.6** Surface tension components for solid surfaces

Surface	$\gamma$ (mN m <sup>-1</sup> )	$\gamma^d$ (mN m <sup>-1</sup> )	$\gamma^p$ (mN m <sup>-1</sup> )
PTFE	18–22	18–20	0–2
Polyethylene	33	33	0
PMMA	41	30	11
PET	44	33	11

combines the contributions from two components A and B:

$$\gamma_{AB} = \gamma_A + \gamma_B - 2(\gamma_A^d \gamma_B^d)^{1/2} - 2(\gamma_A^p \gamma_B^p)^{1/2} \quad (9.12)$$

Thus the surface tension components of liquids can be obtained from measurement of interfacial tension against an immiscible probe. Other semi-empirical expressions exist for relating the polar and dispersive components of a liquid and solid with the contact angle. These models are based on methods of combining Hamaker constants such as the geometric mean theorem for instance [7]. The Owens–Wendt model, Equation (9.13), is one example:

$$1 + \cos\theta = 2(\gamma_s^d)^{1/2} \left[ \frac{(\gamma_l^d)^{1/2}}{\gamma_{lv}} \right] + 2(\gamma_s^p)^{1/2} \left[ \frac{(\gamma_l^p)^{1/2}}{\gamma_{lv}} \right] \quad (9.13)$$

This expression can be used to estimate the solid surface energy and the contributions from both the dispersive and polar components. The most simple method is to measure contact angles with two probe liquids, one polar and one non-polar, but whose individual components are known. The simultaneous equations can be solved to give estimates of the solid surface properties.

Table 9.6 shows the numerical values obtained for the dispersive and polar components from a series of materials, and there is a reasonable agreement. PTFE, for instance, which is a very non-polar material, is more or less completely dispersive. Polyethylene is non-polar and again the only contributing component is dispersive. For polymethylmethacrylate (PMMA) there is a contribution from a polar component, as is the case for polyethylene terephthalate (PET)[9].

## 9.15 Wettability Envelopes

It may be imagined that up until now we have a *rule* which tells us if wetting will occur when placing a liquid on a solid. This comes from knowing the surface tension of the liquid and the surface energy of the solid. Thus the rule which has been formulated states that ‘a liquid, having a lower surface tension than the solid surface energy, will wet that solid’. In practice it is found that this is not always the case and thus it is not an immutable rule.

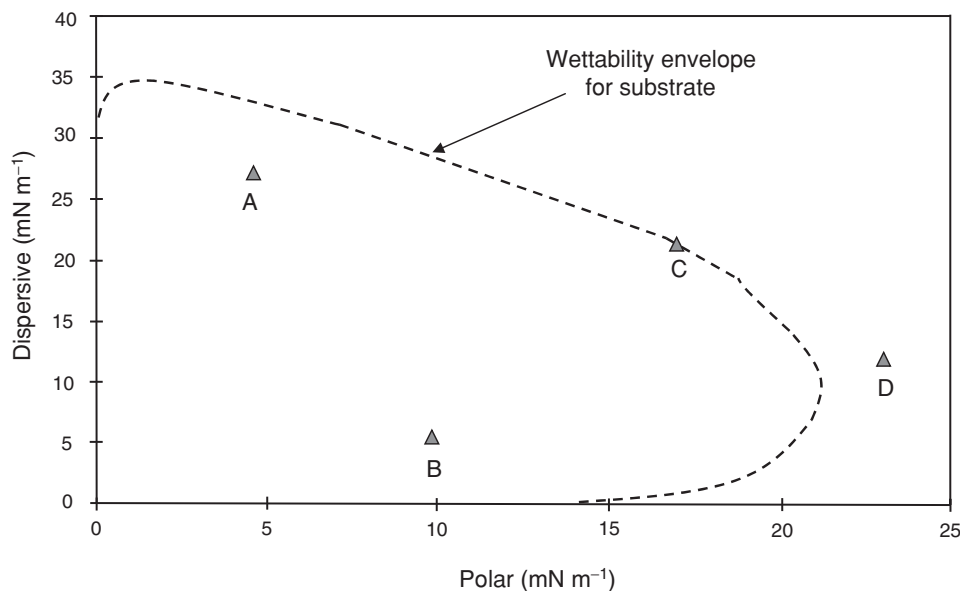
A 2D map of wetting can be constructed by using the components of surface tension and a plot produced which is designed to show where wetting will occur. To illustrate this the Owens–Wendt model described previously has been used to construct a plot.

It is suggested above that there are other ways of understanding how contact angles, hence degrees of wetting, arise from an understanding of the forces existing in the materials and

between the materials. Rather than considering surface tension (or surface energy) as a single component it can be seen that surface tension is the sum of individual components, dispersive and polar components for instance, and these can be summed to yield the surface tension. It is possible to take these components which have been generated by the empirical expressions and draw (2D) maps of wetting.

The experimental programme required to produce the necessary information is relatively simple and follows the rationale described in Section 9.14. The contact angles for a solution on two standard substrates, for instance glass and polyvinylchloride (PVC), are determined. One of the surfaces is polar and the other is non-polar. Following this the contact angles of two standard liquids on the substrates are determined (iodomethane and water, for instance). The contact angle on the substrate of the liquid of particular interest is subsequently determined. From this information the polar and dispersive components are derived by using the Owen–Went equation. For the unknown sample these components are plotted against each other. An example plot is shown in Figure 9.13.

The dispersive component is plotted along the  $y$  axis against the polar component along the  $x$  axis. Four points labelled A, B, C and D are shown on the plot. An envelope, the ‘wettability envelope’, is also plotted. The envelope is created when the Owen–Went model is solved for the case of a contact angle of  $90^\circ$ ; so the area bounded by the axes and the curve is less than  $90^\circ$  and that outside this boundary is greater than  $90^\circ$ . For each of the four liquids it can be seen that A and B will wet the substrate. From this it is clear that the two materials A and B, which have different overall surface tensions, can be plotted on a map, and both can be seen to be wetting. The values of polar and dispersive contributions have been calculated previously. This starts to give an understanding of wettability. It is also clear that D is dewetted, since it sits outside that envelope. In the case of liquid C it has a contact angle of



**Figure 9.13** Wettability envelope for a substrate with a series of liquids.

90° and so is on the border of wetting and dewetting. It is possible to make a comparison here between this type of approach and the calculation of solubility parameters. The Hildebrandt solubility parameter [3] gives a single value of solubility, whereas it is also possible to break the solubility parameter into components, as is done in the Hansen solubility parameters [3, 10], or partial solubility parameters. In the latter a three-component coordinate set of parameters can be identified, which map out the solubility in a more detailed manner and give a greater insight into the solubility of materials, polymers for instance, in different solvents. Thus the solubility of a material is put onto a 3D map which has axes of polar, dispersive and hydrogen bonding. The same principle is adopted here, but using surface tensions, to see how wettability is influenced by the components of surface tension. This gives another view and method of predicting the wetting behaviour of substrates with liquids.

## 9.16 Measurement Methods

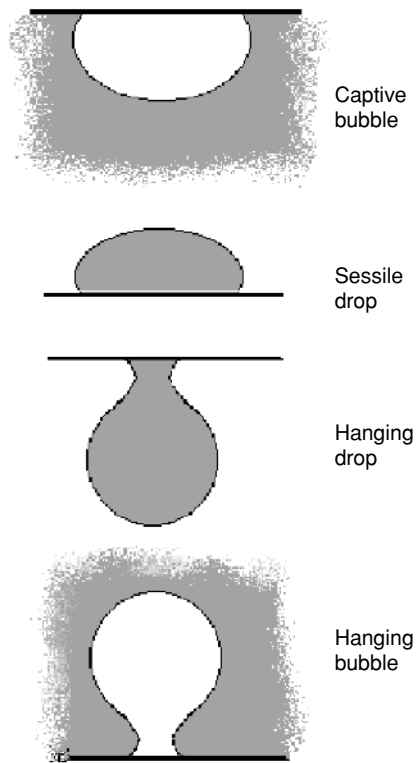
Clearly it is of major importance to be able to make measurements of contact angles, interfacial tensions and surface tensions, and this itself is quite a broad subject and topic to consider. What is presented below is an identification of some of the more important techniques. It is suggested that if the reader requires a more in-depth description further reading can easily be found on each of the techniques [1].

Capillary rise experiments can be used for both surface tension and contact angle. In each case the experimental elegance employed can potentially offer a high degree of accuracy producing contact angles with  $\pm 0.1^\circ$  accuracy. This is fine in an academic sense; however, new commercially available equipment gives rapid capabilities for industrial applications. Contact angle can be routinely measured by imaging techniques, essentially computer frame grabbing. Du Nouy ring still offers a reasonable surface tension measurement method but additionally is still the most appropriate method for interfacial tensions. Spinning drop again is an interfacial tension measurement technique but is useful only for ultralow interfacial tensions. Maximum bubble pressure techniques are able to study the development of surface tension with time, as at small times a bubble blown in surfactant solution will give a value close to the pure solvent, whilst at long times the equilibrium value will be given.

Techniques requiring drop or bubble shape analysis have been advanced over recent years because of greatly improved computational methods when applied to an image of the drop or bubble profile. The method, known as ‘axisymmetric drop shape analysis’ [11], can be performed on the measured profiles of captive bubbles, sessile drops, hanging drops and hanging bubbles, as shown in Figure 9.14. The principle of the technique is that the pressure drop across a curved surface given by the Laplace equation is in balance with the gravitational pull on the bubble or the drop as shown by

$$\gamma \left( \frac{1}{R_1} + \frac{1}{R_2} \right) = \Delta p_0 + \rho g z \quad (9.14)$$

From this we can develop some idea of the drop shape from the balance between the Laplace pressure and the gravitational force, since it is a function of surface tension of the interface,  $\gamma$ . The appropriate expressions are written in terms of surface tension, the principal radii of the bubble or the drop,  $R_1$  and  $R_2$  and the pressures associated with them,  $\Delta p$ , the pressure drop



**Figure 9.14** Axisymmetric drop shape analysis measures the profiles of symmetrical drops and bubbles.

across the interface. This has two contributions:  $\Delta p_0$  the pressure difference at a reference plane and an additional pressure which arises from the gravitational component at any point on the surface. This is given by the density difference ( $\rho$ ) between the bubble or drop and the continuous phase and its vertical distance  $z$  from the reference plane. The reference plane is a datum line from which values of  $z$  can be established, hence its position needs to be established and a calibration taken to ensure determining accurate values of  $z$ . The principal radii can be used to calculate the surface tensions from knowledge of the density differences, the height to  $z$  above the reference plane and the curvature at that height. The analysis involves a complex set of first order differential equations that are solved numerically to give the results. From the measured profiles it is possible to obtain parameters such as surface tension, contact angle, the drop radius, drop volume and surface area.

The method itself is not new and goes back very many years to a set of tabulated calculation results of Bashforth and Adams [12]. The functions are look-up tables which make corrections to the observed values of diameter, height and radius of the drop or bubble. Using these values the surface tensions can be calculated.

There are a number of commercially available machines and because of the nature of the experiment it is rapid, simple and produces a wealth of information. Because the method observes the drops or bubbles optically the technique requires the magnification to be

calibrated so that absolute dimensions are known. This can be done in a number of ways but the preferred method is to measure an object of known size, e.g. a sapphire ball. It is worth noting that the methods require an axially symmetric drop or bubble. This condition would be expected to be fulfilled for all the cases with the possible exception of the sessile drop. It is often noticed that a drop spreading on a surface does so in a series of jumps. These we can argue are due to a number of causes, surface roughness, oil or contaminant on the surface, different exposed surface energy domains, for instance. The consequence of this is that the drop may not be axially symmetric. Therefore, a degree of care must be exercised in performing these experiments.

An additional advantage of using this type of experimental equipment is that previously obtained images can be presented to the analysis system. Hence, photographs of high temperature molten metals in contact with a surface, for instance, can be analysed to give information regarding the surface tension and contact angle.

## 9.17 Conclusions

The topic of wetting as outlined in this chapter is seen to be a fascinating study. Only a glimpse of the complete topic is shown here, but as more advanced reading into the subject is undertaken, it will be appreciated that there is still much to understand. Many of the expressions and arguments shown in the text can at best be described as semi-empirical. This confirms our belief that wetting is still a developing science. Perhaps it is because the subject has such immense utility to industrial processes that our full understanding has lagged behind the experimental development. Naturally, within the text here there are many missing components of the subject, and one which deserves further reading is that of the dynamics of wetting [13]. This is a rich topic for investigation with tremendous implications on industrial processing.

## References

1. Adamson, A.W. and Gast, A.P. (1997) *Physical Chemistry of Surfaces*, 6<sup>th</sup> edition. Wiley, New York.
2. Young, T. (1805) *Phil. Trans. R. Soc. (London)*, **95**, 65.
3. Barton, A.F.M (1991) *CRC Handbook of Solubility Parameters and Other Cohesion Parameters*, 2<sup>nd</sup> edition. CRC Press, Boca Raton, FL, Chapters 1 and 2.
4. Zisman, W.A. (1964) *Adv. Chem. Ser.*, A **43**, 1.
5. Garret, P.R. (ed.) (1993) *Defoaming, Theory and Industrial Applications*, Surfactant Science Series vol. 45. Dekker, New York.
6. Schulte, H.G. and Hofer, R. (2003) In D.R Karsa (ed.), *Surfactants in Polymers, Coatings, Inks and Adhesives*. Applied Surfactant Series vol. 1. Blackwell, Oxford.
7. Girifalco, L.A. and Good, R.J. (1957) *J. Phys. Chem.*, **61**  
Good, R.J. (1964) *Adv. Chem. Ser.*, **43**, 74.
8. Fowkes, F.M. (1962) *J. Phys. Chem.*, **66**, 382  
Fowkes, F.M. (1964) *Adv. Chem. Ser.*, **43**, 99.
9. Van Krevelen, D.W. (1990) *Properties of Polymers, Their Correlation with Chemical Structure; Their Numerical Estimation and Prediction from Additive Group Contributions*, 3<sup>rd</sup> edition. Elsevier, Amsterdam.



- 
10. Patton, T.C. (1979) *Paint Flow and Pigment Dispersion*, 2<sup>nd</sup> edition. Wiley-Interscience, New York.
  11. Neuman, A.W. and Spelt, J.K. (eds) (1996) *Applied Surface Thermodynamics*, Surfactant Science Series vol. 63. Dekker, New York.
  12. Bashforth, F. and Adams, J.C. (1883) *An Attempt to Test the Theories of Capillary Action*. Cambridge University Press, Cambridge.
  13. Blake, T.D. (1993) In J.C. Berg (ed.), *Wettability*. Dekker, New York.

# Chapter 10

## Aerosols

*Jonathan Reid*

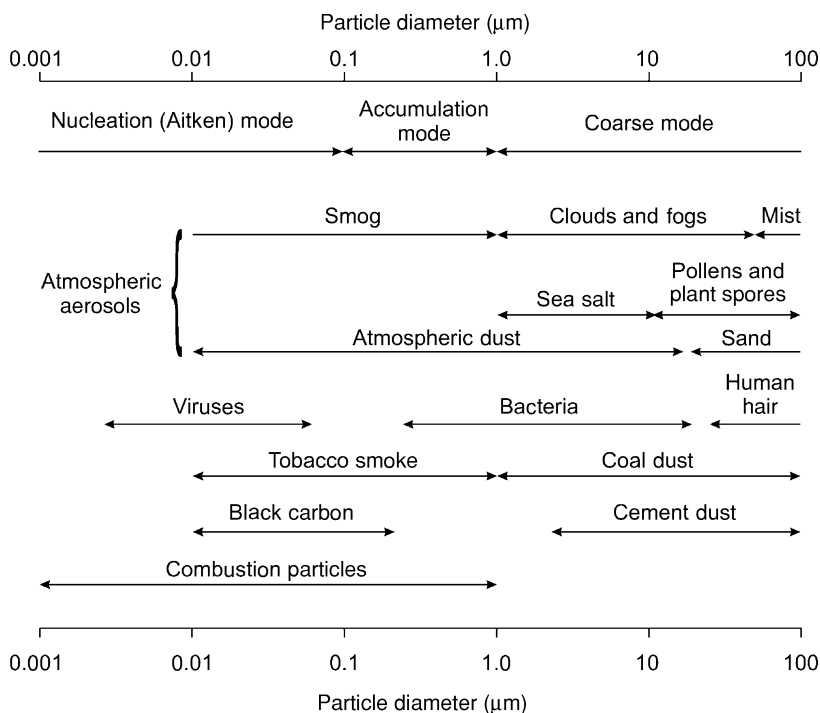
### 10.1 Introduction

An aerosol can be defined as a mixed phase system that arises when solid or liquid particles are dispersed in a gas. The particle phase represents a very small fraction of the aerosol volume, and the bulk properties of the aerosol resemble very closely the bulk properties of the gas phase. Particle diameters can span five orders of magnitude, ranging from nuclei at the nanometre scale to cloud droplets, dust particles and spray at the 100  $\mu\text{m}$  scale. This corresponds to a variation in volume and mass of 15 orders of magnitude and five orders of magnitude in surface-to-volume ratio, bridging the divide between the microscopic behaviour of molecules and the macroscopic properties of materials.

As a result of the wide variation in size, concentration, composition, morphology and phase, a detailed characterisation of the state of an aerosol is challenging [1–3]. Ideally, all of the physical and chemical properties of a particle would be measured, examining, for example, the variation in particle composition with particle size. To complicate aerosol analysis further, the highly dynamic nature of aerosol particles can lead to pronounced and rapid temporal variations in particle size and composition, arising from the large surface area interacting with the surrounding gas phase. Clearly, compromises must be made in the analysis of aerosols. Particles may be co-analysed within certain broad size ranges. It may not be possible to measure aerosol composition in real time, but only off-line using sensitive analytical tools in the laboratory. In some circumstances it may only be possible to measure the aerosol mass within a standard gas phase volume, obtaining no information on particle size distribution.

Recent epidemiological studies have highlighted the significant impact that aerosols can have on human health, increasing mortality and morbidity at lower particle concentrations than had previously been thought. Although the chemical composition is important in determining toxicity with particles acting to adsorb and concentrate potentially carcinogenic species, physical properties are also important. Toxicity can be correlated with particle size, shape, electrical charge and solubility, with fine particles ( $<1\ \mu\text{m}$  in diameter) able to penetrate deep into the respiratory tract, increasing the rate of respiratory, pulmonary and cardiovascular diseases.

Not only can pollutant particles be hazardous for health, but they also play a major role in determining the chemical balance of the atmosphere. The origins of acid rain and the ozone hole are but two examples. In addition, aerosols impair visibility and have a significant



**Figure 10.1** Classification of particles according to size and examples of common aerosols.

influence on the radiative balance of the atmosphere, although their contribution is as yet unquantified.

Particles are commonly classified according to their size (Figure 10.1). Particles smaller than  $0.1 \mu\text{m}$  are referred to as nucleation particles (or Aitken nuclei) and are generated by gas-to-particle conversion. Growth of the nucleation mode by condensation and coagulation leads to the generation of the accumulation mode with particles ranging from  $100 \text{ nm}$  to  $1 \mu\text{m}$  in size. Larger particles are referred to as the coarse mode. Particles generated by incomplete combustion processes can range from a few nanometres in size to  $1 \mu\text{m}$ , spanning the nucleation and accumulation modes. Pollens, dust and sea salt particles are mostly coarse particles larger than  $1 \mu\text{m}$ .

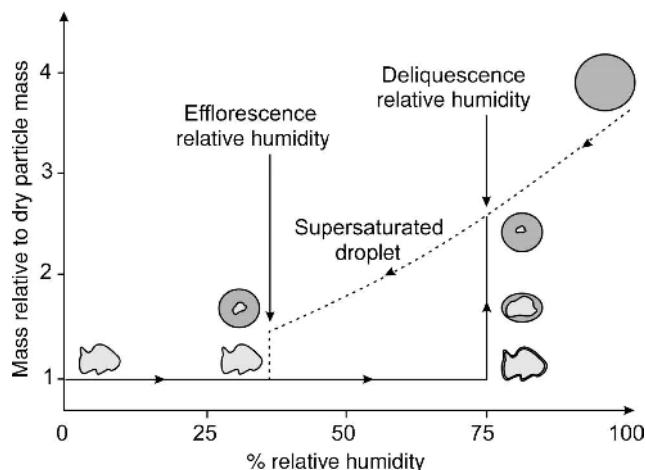
Particle concentrations are usually expressed as the number of particles per cubic centimetre ( $\text{particles cm}^{-3}$ ) or a mass per cubic centimetre (usually  $\mu\text{g cm}^{-3}$ ). Typical particle concentrations in the atmosphere are  $10\text{--}1000 \text{ cm}^{-3}$  for particles smaller than  $1 \mu\text{m}$ , falling to  $<1 \text{ cm}^{-3}$  for larger particles. Particle concentration and size distribution are usually indicative of source. Particle concentrations are usually lowest over the oceans and largest in urban areas with typical concentrations of  $100 \text{ cm}^{-3}$  and  $>1 \times 10^5 \text{ cm}^{-3}$ , respectively. Although the maximum particle concentrations usually occur in the accumulation mode, these particles are comparatively small in volume when compared with the coarse mode. Thus, most of the aerosol volume ( $\mu\text{m}^3 \text{ cm}^{-3}$ ) and mass occurs in the coarse mode. However, the larger surface to volume ratio of smaller particles and their high number densities

lead to a surface distribution ( $\mu\text{m}^2 \text{cm}^{-3}$ ) that peaks in the accumulation mode. Particles in the accumulation mode dominate aerosol chemistry in the atmosphere.

Aerosol composition varies with particle size, source and lifetime. Indeed, individual particles may not be homogeneous in composition, but may be multi-phase and multi-component. Solid inclusions may be contained within a liquid host. Further, the composition of the particle surface may be radically different from the composition of the aerosol bulk. Dramatic changes can occur in particle size and composition with variation in relative humidity. A combination of all of these complex factors illustrates the problems associated with characterising the chemical composition of atmospheric aerosols.

Aerosols may be emitted directly into the atmosphere or may be generated by physical and chemical processes within the atmosphere. These are known as primary and secondary aerosols, respectively. Primary aerosols include dust particles and sea sprays. Reaction of gases or condensation onto pre-existing aerosol produces secondary aerosol. Natural sources are estimated to deliver  $3100 \text{ Tg yr}^{-1}$  (Tera-grams per year) into the atmosphere, while anthropogenic sources are estimated to contribute  $\sim 450 \text{ Tg yr}^{-1}$ . It should be remembered, however, that anthropogenic sources are spatially very concentrated. Natural aerosols arise from desert sands, rock weathering, soil erosion and volcanoes, and their chemical constituents include aluminosilicates, ores, clays, organic matter and soots produced in biomass burning. The total carbon loading present in atmospheric aerosols is typically divided into an organic carbon (OC) fraction and an elemental carbon (EC) fraction. Carbon from carbonates is estimated to contribute less than 5% to the total particulate carbon loading. Marine aerosol is composed of metal halides, sulfates and organic compounds. Anthropogenic sources include chemical emissions, fossil fuel combustion and heavy industry, with particulate emissions containing soot, fuels and heavy metals.

The transformation and ageing of aerosols can occur by condensational growth or evaporative loss, resulting in a change in partitioning between the gas and condensed phases. Growth occurs in liquid aerosols when the vapour pressure of the condensed phase component is lower than the partial pressure of the component in the gas phase. The mass growth rate is proportional to the surface area of the droplet. The state of a soluble solid inorganic salt is dependent on the relative humidity (RH) of the environment, as shown in Figure 10.2. The RH is defined as the ratio of the water vapour density to the saturation water vapour density, expressed as a percentage. At low RHs, the salt exists as a solid particle with a certain dry particle diameter. At an RH known as the deliquescence point, the anhydrous particle grows, forming a hydrated salt solution. The deliquescence point for ammonium sulfate occurs at 75% RH. For solid multi-component aerosols, the solubility of each component must be known to predict the partitioning of the chemical species between the solid and aqueous phases. At RHs above the deliquescence point, the particle continues to grow, stabilising at a size governed by the dry particle diameter and the RH. The variation in equilibrium particle size with RH is described by Köhler theory, which reflects the balance between the Kelvin effect that tends to increase the vapour pressure of the aqueous component and the solute effect that tends to reduce it [4]. If the RH decreases, the liquid droplet loses water by evaporation, remaining as a metastable solution at RHs lower than the deliquescence point. Once the efflorescence point is reached, a solid nucleus forms and crystallisation occurs rapidly.

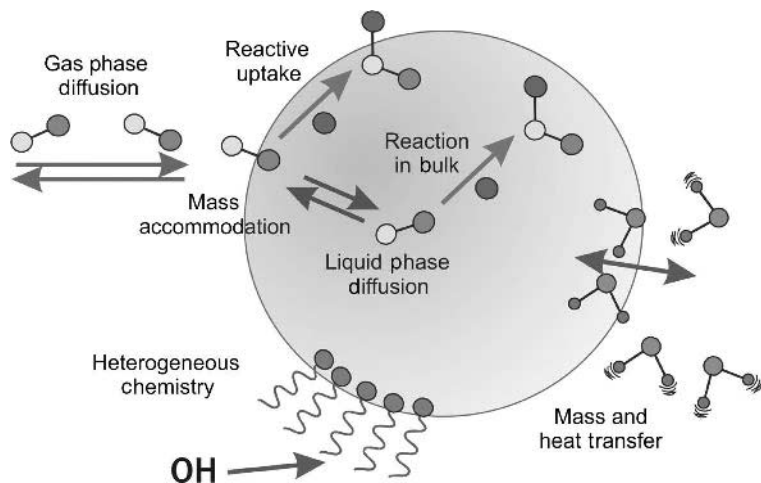


**Figure 10.2** The variation in particle composition of a soluble inorganic salt particle with relative humidity (Reid and Sayer, *Chemical Society Reviews*, 32 (2003) p. 71 – reproduced by permission of The Royal Society of Chemistry).

Particle transformation can also occur by coagulation through particle collision. The high diffusion constants of nucleation mode particles lead to rapid diffusional loss through coagulation. The large cross sections of particles larger than  $10\ \mu\text{m}$  leads to an appreciable growth rate through coagulation for such large particles.

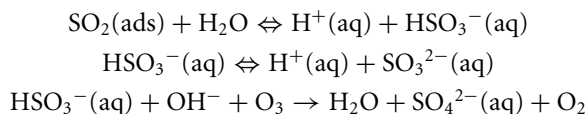
Deposition provides the dominant loss mechanism for most aerosols through impaction, diffusion or gravitational settling. Small particles are buoyant and are lifted by upwelling air currents. However, large particles reach a terminal velocity, and gravitational settling leads to aerosol loss. A gradient in particle concentration can be established, with lower particle concentrations near surfaces due to deposition. Brownian motion leads to diffusion of the remaining particles towards the surface followed by deposition. Impaction occurs for particles of high mass and inertia that are not able to adapt their path with the deflected air flow around an object, leading to impaction. In the atmosphere, wet deposition is also an important loss mechanism for particles  $0.1\text{--}10\ \mu\text{m}$  in diameter through incorporation in existing rain droplets and wash out.

The chemical transformation of particles can occur not only through coagulation and mixing, but through heterogeneous chemistry with trace reactant species. At the microscopic level, reactant molecules must first diffuse through the gas phase, as shown in Figure 10.3. It is useful to define a relative length scale, known as the Knudsen number ( $Kn$ ), which is equal to the ratio of the mean free path of the diffusing molecule to the particle radius [4]. At low Knudsen numbers ( $\ll 1$ , known as the continuum regime), the mean free path is much smaller than the size of the particle and the gas phase behaves as a continuous fluid surrounding a large particle. If the Knudsen number is large ( $\gg 1$ , known as the free molecule regime), the particle is much smaller than the mean free path of the gas phase molecules and the gas phase molecules move discretely around it. On collision with the surface of the particle, the reactant gas molecule must accommodate on the surface. Subsequently reaction can occur either on the surface of the particle or in the particle bulk if it is a liquid aerosol. Heterogeneous chemistry competes with chemistry that occurs in the gas phase as a result



**Figure 10.3** Schematic of the fundamental processes involved in the partitioning of gas phase species to the aerosol condensed phase.

of the lower activation barriers that reactants can encounter in the condensed phase due to solvation of ionic intermediates, reactants and products. As an example, the oxidation of  $\text{SO}_2$  to sulfuric acid occurs rapidly in humid air containing aerosol particles leading to the generation of acid rain. This is a consequence of the lower activation energies that can occur in the condensed phase for the ionic reaction mechanism, compared to the much larger activation energies for reaction in the gas phase:



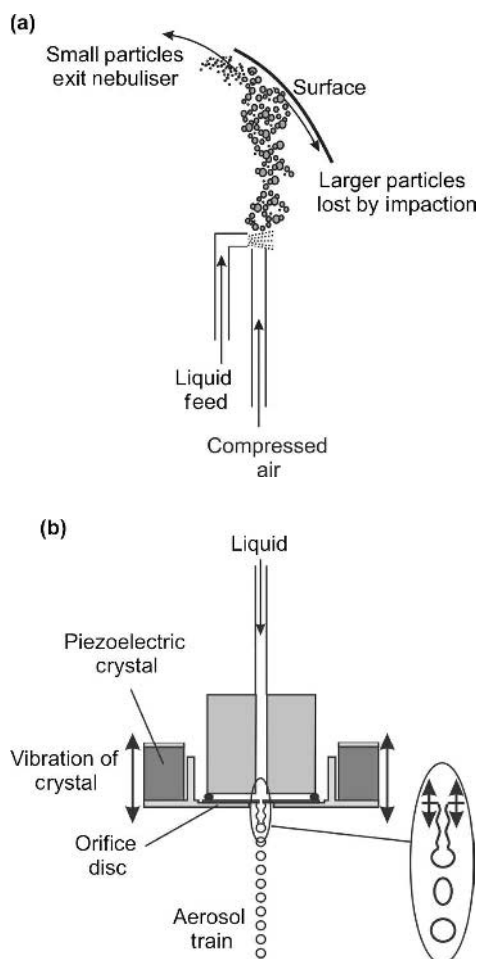
In the remainder of this chapter, we focus on the techniques available for generating and sampling aerosols, characterising size distributions and determining chemical composition.

## 10.2 Generating and Sampling Aerosol

Generating test aerosols with well-defined size distributions and compositions is crucial for calibrating and testing many of the characterisation techniques that are described below. In some instances, monodisperse aerosol of known size, shape and density may be required to test sampling systems and to calibrate particle sizing instruments. The aerosol generator must yield a reproducible output of uncharged spherical particles of constant concentration and size. Polydisperse aerosol may be needed to test analytical instruments under controlled laboratory conditions. In addition, a sampling system that collects a representative sample of an aerosol is crucial if an accurate analysis of the particle size distribution and concentration is to be achieved. Aerosol generation and sampling will now be considered.

### 10.2.1 Generating aerosol

A variety of techniques exist for atomising liquids. A compressed air nebuliser is the simplest method for generating liquid aerosol droplets, as illustrated in Figure 10.4. Compressed air at a pressure of 30–250 kPa exits at high velocity from a nozzle or orifice. The low pressure resulting from the Bernoulli effect in the exit region of the nozzle draws liquid from a liquid reservoir into the gas stream and breaks the liquid into droplets [3]. The resulting aerosol spray is directed onto a surface where large particles are lost by impaction and small particles remain in the gas flow. A polydisperse stream of droplets with mass median diameters of 1–5  $\mu\text{m}$  results with concentrations as high as  $10^7$  particles  $\text{cm}^{-3}$ . If liquids of low volatility are nebulised, the size distribution may remain stable for many minutes. If a non-volatile salt is dissolved in a volatile solvent, evaporation of the volatile solvent may lead to the formation of solid particles with a final diameter that is determined by the initial volume of the liquid



**Figure 10.4** (a) A compressed air nebuliser and (b) A vibrating orifice aerosol generator (VOAG).

droplet and the volume fraction of the solid material. In much the same way, nebulisation of mixtures of highly volatile and low volatility liquids can lead to small liquid droplets.

Ultrasonic nebulisers operate by focusing ultrasonic waves, generated with a piezoelectric crystal, onto a liquid surface. Capillary waves form at the surface and these break to form a dense aerosol cloud, with the droplet size related to the modulation frequency of the piezoelectric crystal. Ultrasonic nebulisers are able to produce a higher volume of aerosol at higher concentration than compressed air nebulisers.

Nebulising a liquid suspension of solid particles can lead to a controlled source of monodisperse solid particles. Monodisperse homogeneous polystyrene and polyvinyl toluene latex spheres, 0.01–30  $\mu\text{m}$  in diameter, can be purchased with standard deviations in size of only a few per cent. Once dispersed in a solvent and atomised, the solvent evaporates leaving monodisperse spheres for calibrating instruments.

For droplets larger than 5  $\mu\text{m}$  in diameter, a vibrating orifice aerosol generator presents an alternative [3]. A liquid jet is unstable to mechanical disturbance. By applying a regular mechanical vibration to a liquid jet formed by forcing liquid through a small orifice, the jet can be forced to break up into droplets with a monodisperse size, determined by the frequency of the modulating vibration and the size of the orifice. The mechanical vibration is provided by a piezoelectric crystal. Initial droplet diameters from 5 to 200  $\mu\text{m}$  can be readily achieved with concentrations of 10–500  $\text{cm}^{-3}$ , although coagulation can cause rapid increases in particle size unless the aerosol is dispersed in a flow of air. Solid particles can be generated by dissolving a solid non-volatile salt in a volatile solvent.

The simplest method for dispersion of a dry powder is to use a gravity feed of powder into a high-velocity air stream. Concentrations of up to 100  $\text{g m}^{-3}$  can be achieved, although this is dependent on the particle size range and shape, and on the moisture content of the powder. Particle sizes between <1 and 100  $\mu\text{m}$  can be dispersed. Dry and hydrophobic powders disperse more readily than moist and hydrophilic powders. Agglomeration in the powder to be dispersed can be a major problem. This can be overcome by introducing the powder into a fluidised bed consisting of 200  $\mu\text{m}$  diameter beads. Particle charging during the dispersion process can reduce the dispersion efficiency, giving rise to variations in the concentration of the dispersed aerosol.

### 10.2.2 Sampling aerosol

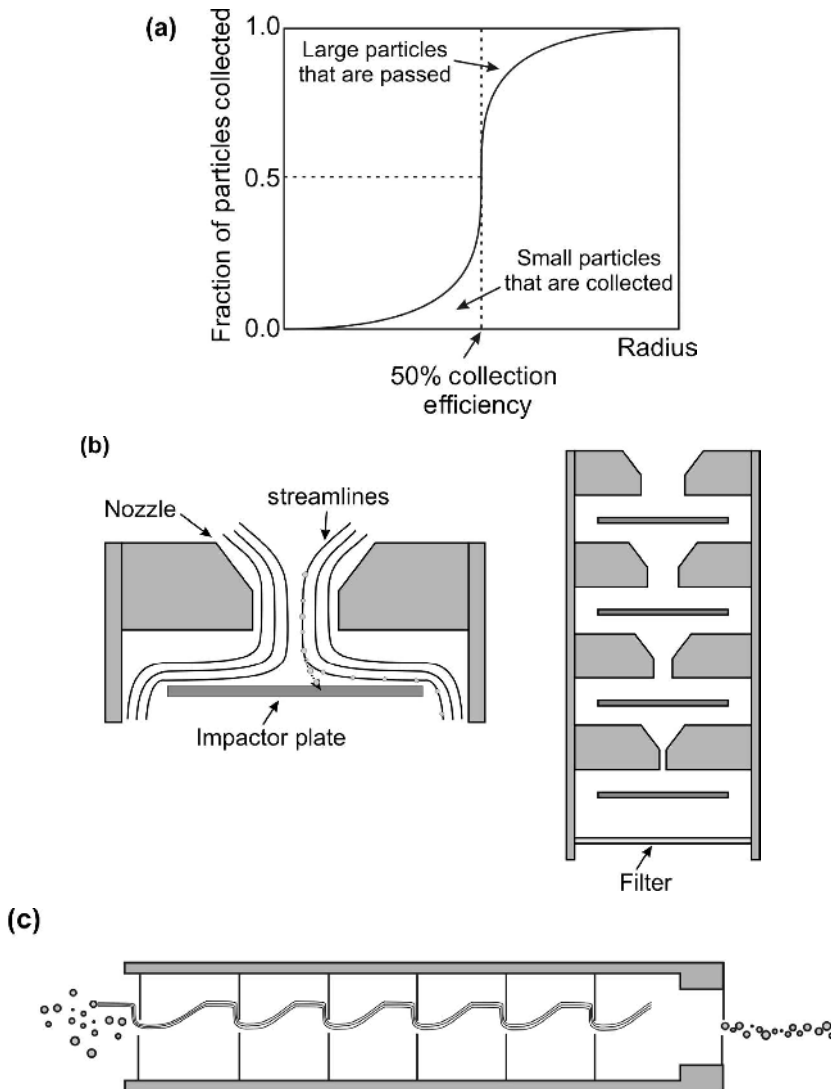
To avoid particle loss in the region of a sampling inlet, irrespective of particle size or inertia, isokinetic sampling must be achieved. The sampling inlet must be aligned parallel to the gas flow and the gas velocity in the inlet must be identical to the velocity of the gas approaching the inlet. If these are not achieved, distortions in size distributions passing through the inlet may result. In addition, any change in the gas phase environment during sampling can lead to dramatic changes in particle size and composition. Sampling particles through an inlet into an instrument at reduced pressure can have a significant impact on the partitioning of volatile components between the condensed and gas phases. Loss of a volatile component, such as an organic component or water, can lead to changes not only in composition, but also in size and even particle phase.

Following sampling of an ambient aerosol, it may be advantageous to select a specific size range for analysis in a controlled manner. Inlets can be designed to deliver a sharp cut in the size distribution, passing particles smaller than a certain aerodynamic diameter, as illustrated



in Figure 10.5a. Particles are classified according to their inertia by impactors and cyclones (Figure 10.5b) [3, 5]. Particles are accelerated through a nozzle or circular jet towards a substrate positioned at  $90^\circ$  to the gas flow in a conventional impactor. Particles larger than a limiting size have sufficient inertia to cross the flow streamlines and impact on the substrate, while smaller particles follow the deflected streamlines avoiding impaction on the substrate. Impactors can be designed with different size-dependent collection efficiencies.

When a number of impactors are operated in series with progressively smaller nozzle sizes and hence size cut points, a wide range of particle sizes can be fractionated and collected. This is referred to as a cascade impactor. Impactors can allow an accurate determination



**Figure 10.5** (a) Inlets are designed to collect particles larger than a specified size with a sharp cut point, (b) an impactor and cascade impactor and (c) an aerodynamic lens.

of particle concentrations for particle sizes smaller than  $10\text{ }\mu\text{m}$ , and typical fractionations include the separation of 10, 2.5 and  $1\text{ }\mu\text{m}$  particles. A low pressure cascade impactor can enable the selection of particles in the nanoparticle range, and particle sizes down to 50 nm can now be routinely studied. A fraction of particles may bounce on encountering the substrate, particularly if it is solid. This can be minimised by coating the substrate with oil or grease. Alternatively, this problem may be avoided entirely by replacing the substrate with a receiving tube. This is known as a virtual impactor.

Cyclones are inexpensive and easy to maintain, although their performance must be determined by calibration. The inlet geometry is such that sampled air enters tangentially into a cylindrical or cone shaped chamber, travels around the axis of the cyclone several times and exits to a filter at the top of the cyclone. Small particles are sampled by the filter, while larger particles are deposited on the walls by the centrifugal force and may be collected at the bottom of the cyclone.

A particle beam with a small divergence can be used to achieve high transport efficiencies from a sample region to a detection region. An air stream can be introduced upstream of a sampling nozzle to assist in the confinement of particles to a narrow region near the flow axis. The particles retain the motion they have attained during their acceleration through the nozzle due to their large inertia. However, sheath air dilutes the particles and requires additional gas handling systems. Alternatively, a collimated flow of particles can be achieved by passing the aerosol through a series of axisymmetric contractions and enlargements before the nozzle expansion, known as an aerodynamic lens (Figure 10.5c). Particles smaller than a critical size can be confined very closely to the axis of the flow, with particle beams confined to waists as little as a few millimetres.

Aerosol samples are commonly acquired on fibrous or porous membrane filters. Particle removal occurs as a result of collision and adsorption of the particle to the fibre surface through interception, inertial impaction, diffusion or electrostatic interaction [6]. Fibre porosities are generally very high, even as high as 99%, but collection efficiencies are typically also greater than 90%. Fibres range from  $<1\text{ }\mu\text{m}$  to  $100\text{ }\mu\text{m}$  in diameter in fibrous filters, forming an entangled mat in which the fibres are arranged perpendicular to the air flow. Fibrous filters are typically composed of glass fibres, plastic fibres or cellulose fibres. Porous membrane filters are lower in porosity, with particles depositing on the inside of pores,  $0.01\text{--}10\text{ }\mu\text{m}$  in diameter, as the gas flow follows a complex path through the pores forming the membrane. Examples of materials used for porous membrane filters include Teflon, polyvinyl chloride and sintered metals. Thin polycarbonate films can be etched to produce capillary pore membrane filters, which consist of an array of uniform cylindrical pores  $0.03\text{--}12\text{ }\mu\text{m}$  in diameter. Nucleopore filters with pores  $5\text{--}12\text{ }\mu\text{m}$  in diameter can be used to remove coarse particles.

### 10.3 Determining Particle Concentration and Size

To characterise the aerosol dispersion, it is important to determine the number concentration of particles, the mass concentration and size distribution. The strengths and deficiencies of some of the common methods will now be described. Other properties, such as composition and morphology, will be discussed later.

### 10.3.1 *Determining number concentration*

The number concentration provides a measure of the number of particles per unit volume. The most commonly applied instrument is a condensation particle counter (CPC), also known as a condensation nucleus counter. Particles undergo rapid growth in a supersaturated environment to a size at which they can be detected by light scattering. Individual particles  $<10$  nm in diameter can be detected. The lower limit of detectable size is determined by transport efficiency and the degree of supersaturation achieved during the growth phase. However, with growths in particle diameter as much as a factor of 1000, the original particle represents a volume fraction as low as  $10^{-9}$  of the final detected droplet. To detect particles  $<10$  nm in diameter, supersaturations of several hundred per cent may be required to achieve particle growth at a sufficient rate for particle detection to be achieved. If the size distribution is broad, particle growth occurs more rapidly for the coarse particles with a large surface area, influencing the sensitivity of the measurement to the fine particles. Commonly used vapours include water and *n*-butyl alcohol. The supersaturation that is achieved is so high that the character of the original particle is irrelevant to the subsequent growth process.

A variant of the CPC is the cloud condensation nuclei counters (CCNCs). The purpose of such an instrument is to assess the concentration of aerosol particles present in the atmosphere that can act as cloud condensation nuclei, undergoing activation to form cloud droplets. Thus, the CCNC is only operated with water vapour at a supersaturation of 0.01–1%.

Individual light scattering events can be used to directly count particles exiting from the growth region or the concentration can be determined indirectly by monitoring the light attenuation. Single particle counting can be achieved at number concentrations lower than  $10^4$  cm $^{-3}$ . While this direct method is limited by uncertainties arising from Poisson counting statistics, the indirect attenuation method is susceptible to inaccuracies in calibration.

### 10.3.2 *Determining mass concentration*

Mass concentration (mass per unit volume) provides an important measure of aerosol loading, particularly for enforcing regulatory standards on air quality. The most common method involves recording the mass loading of filters, under conditions of controlled temperature and relative humidity, prior to and after sampling a known volume of air over a set period. Particles larger than a specific size are removed at the inlet stage. Filters are made from a wide range of materials and can have collection efficiencies exceeding 99%. Particles smaller than  $0.1$   $\mu$ m are collected by diffusion, with the collection efficiency increasing with decreasing size. Particles larger than  $0.5$   $\mu$ m are collected by interception and impaction with collection efficiency increasing with increasing size. Cascade impactors can enable a determination of the cumulative mass distribution within a number of sampling size ranges.

Uncertainties in mass distribution can arise from water absorption/desorption from the filter, evaporative losses of semi-volatile compounds and particle loss during handling. A lower limit on the mass concentration that can be routinely detected is  $2$   $\mu$ g m $^{-3}$ . The attenuation of beta particles with energies of 10–100 keV by a filter can be used to determine the particle loading. In a similar method, particle dosing with a radioactive isotope such as  $^{211}\text{Pb}$

prior to collection on a filter can provide a method for measuring filter loading in an instrument known as an epiphaniometer. The rate of transfer of the radioactive isotope to the condensed phase is proportional to the surface area of the particle for large Knudsen numbers.

Piezoelectric crystals can be used as a sensitive method for determining mass concentrations. The resonant frequency of vibration of the crystal varies with material and with thickness. If the crystal increases in mass, the change in resonant frequency can be used to measure the mass deposited from a known volume of air. A typical sensitivity of  $1000 \text{ Hz } \mu\text{g}^{-1}$  can be achieved for a resonant frequency of 10 MHz. At loadings greater than  $10 \mu\text{g}$ , non-linearities in frequency shift can occur. Measurements of resonant frequency are usually made by reference to an identical piezoelectric crystal held at the same temperature and under condition of the same relative humidity. Such instruments are typically used for mass concentrations in the  $10 \mu\text{g m}^{-3}$ – $10 \text{ mg m}^{-3}$  range.

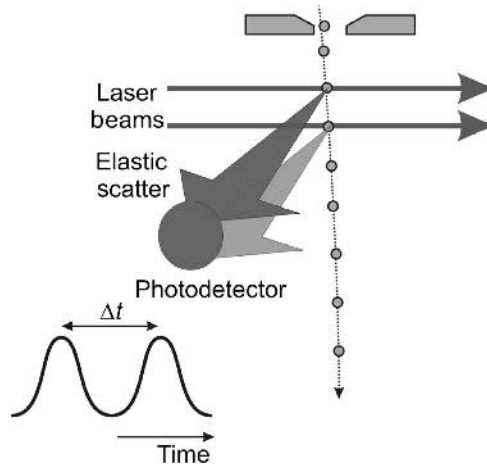
### 10.3.3 Determining particle size

The reported particle size can depend on the technique that is adopted and may be different from the geometric size. Impactors and aerodynamic particle sizers rely on the passage of a particle in a gas flow, and the measured size, referred to as the aerodynamic size, depends on particle shape, density and size. The aerodynamic diameter of a particle is the diameter of the unit density ( $1 \text{ g cm}^{-3}$ ) sphere that has the same settling velocity as the particle. This exceeds the geometric diameter for particles with densities larger than  $1 \text{ g cm}^{-3}$ . For non-spherical particles, it is common to define the Stokes diameter, the diameter of the sphere that has the same density and settling velocity as the particle. More generally for a non-spherical particle, an equivalent diameter is the diameter of the equivalent sphere that would exhibit the same property as that of the non-spherical particle.

Electrostatic methods for determining particle size yield a measure of the electrical mobility of a particle and the electrical mobility size is dependent on particle shape and size but not on density. The intensity of the light scattered from a particle can be used to measure the optical size of the particle, and this is dependent on refractive index, shape and size. The optical size can be very different from the geometric size. Thus, inter-conversion of the different reported sizes is essential, but this is not a trivial procedure. Measurements of particle shape, refractive index and density are essential to perform such a conversion. Methods for determining the aerodynamic, electrical mobility and optical sizes will now be described.

#### (a) Aerodynamic size

Aerosol size distributions can be determined with a cascade impactor, classifying particles according to their aerodynamic diameter. An aerodynamic particle sizer (APS) also exploits the varying inertia of particles with varying size in a gas flow accelerating through a nozzle. The acceleration of a particle through the nozzle increases with decreasing particle size and density. The velocity of a single particle is inferred by determining the time-of-flight of the particle between two probe laser beams, with the scattered light acting to define the timing cycle, as illustrated in Figure 10.6. Although the smallest size that can be probed by the APS



**Figure 10.6** An aerodynamic particle sizer that determines aerodynamic size from the time required to travel between two probe laser beams.

is  $0.2 \mu\text{m}$ , high resolution information on particle size can be achieved in real time. For both of the impaction techniques and the APS, the pressure drop necessary to classify the particle size can lead to a change in relative humidity and particle size, as well as a deformation in particle shape.

The inverse relationship between diffusion constant and particle size is exploited in diffusion batteries. Diffusional loss from a laminar flow to the walls of a conduit results in less than 100% transmission. The transmission of a specific particle size is dependent on the dimensions of the conduit, the gas volume flow rate and the particle diffusion constant. With large diffusion constants, the transmission of smaller particles is reduced through deposition on the walls of the conduit. By measuring the penetration (the ratio of the number of particles leaving to that entering the tube) the diffusion constant and particle size can be determined. Diffusion batteries consist of a tightly packed array of conduits, based on a tight bundle of tubes, a series of closely spaced parallel plates or multiple layers of very fine wire mesh. Particle sizes 2–200 nm in diameter can be determined with a diffusion battery. Measurements of transmission through a number of batteries with different effective lengths, or the same battery at different gas flow rates, can enable the determination of size distributions.

### (b) Electrical mobility size

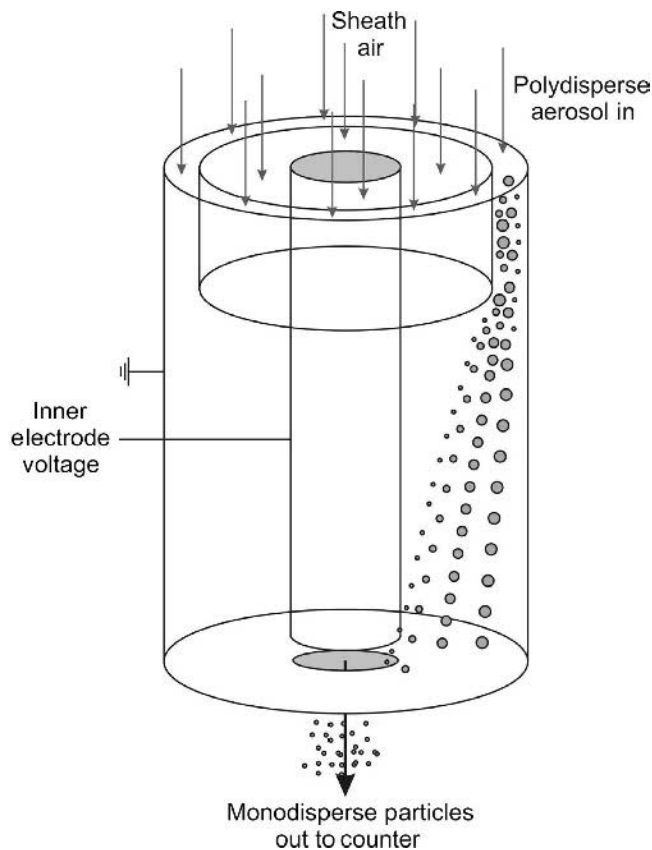
An alternative approach to classifying particles according to their inertia is to classify them according to their electrical mobility,  $Z$ .  $Z$  is defined as the velocity of a particle with a charge  $ne$  in an electric field of unit magnitude:

$$Z = \frac{neC(D_p, p)}{3\pi\eta D_p}$$

$\eta$  is the viscosity coefficient of the gaseous medium,  $D_p$  is the particle diameter and  $C(D_p, p)$  is the slip correction factor. It should be noted that the slip correction factor is dependent

on both the particle diameter and the gas pressure. Electrical mobility size is commonly measured with a differential mobility particle sizer (DMPS), which consists of a differential mobility analyser (DMA, also referred to as an electrostatic classifier) and a particle detector, usually a CPC. The sampled aerosol is exposed to a cloud of positive and negative ions in a bipolar charger, charging the particles to  $\pm 1$ ,  $\pm 2$ , ... charge units with an aerosol mean charge close to zero. The time period spent in the charger is dependent on the gas flow rate; the ion concentration is maintained at a constant value, typically about  $5 \times 10^7 \text{ cm}^{-3}$ . An equilibrium is reached in the bipolar charge distribution, with the fraction of particles showing a positive charge not equal to that showing a negative charge. The contribution of highly charged particles increases with particle diameter; the fraction of 5 nm diameter particles charged to +1 charge unit is  $\sim 2\%$ , while  $\sim 22\%$  of 120 nm diameter particles are charged to +1, 4% to +2 and 0.4% to +3.

The particles pass from the charger into the electrostatic classifier (Figure 10.7), which consists of two concentric metal cylinders, passing into a laminar flow of clean air flowing in the annular space between the two cylinders. In such a coaxial design, the inner cylinder



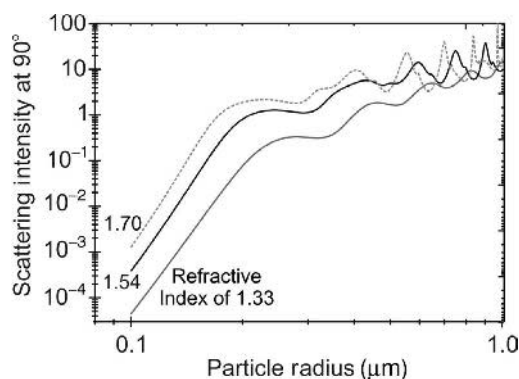
**Figure 10.7** A differential mobility analyser.

acts as the collection rod and is maintained at a negative voltage (0–10 kV) while the outer cylinder is maintained at ground. Particles travel through the electric field between the two cylinders with a well defined trajectory determined by their electrical mobility. Only particles within a narrow mobility range are transmitted through a small slit at the exit of the DMA. The size distribution can be determined by monitoring the number of particles exiting the DMA with variation in classifier voltage. The transfer function determines the probability that a particle with a given electrical mobility is passed through the classifier. The breadth of the transfer function is determined by the flow rate.

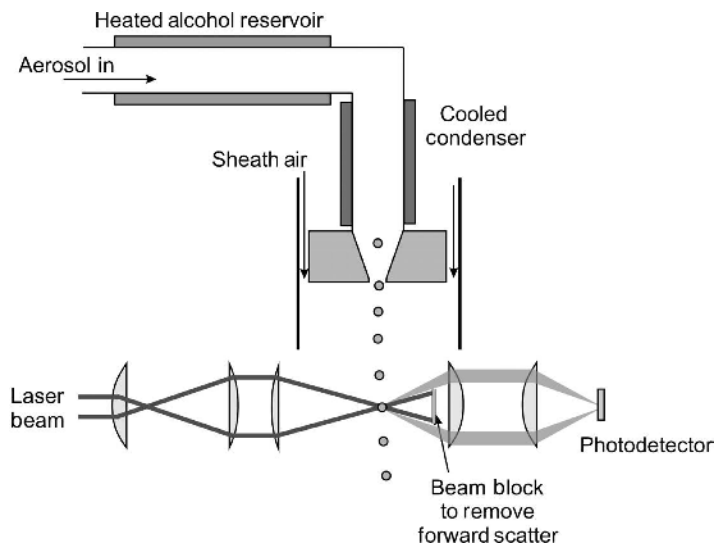
Particle concentrations in the diameter range 3 nm–1  $\mu\text{m}$  can be determined by this approach and timescales for measuring size distributions are typically of the order of 10 minutes. The upper limit imposed on the operating concentration depends on particle size but can be as high as  $2 \times 10^8$  particles  $\text{cm}^{-3}$ . At the lower end of the size range, particle diffusion is significant leading to diffusional losses and broadening of the size distributions. In addition, at a size of 3 nm, only approximately 1% of the particles may be charged and this can lead to a poor detection sensitivity at this low size limit. By scanning the classifier voltage continuously, the time required to measure a size distribution can be significantly reduced to 60 s. Continuous scan is achieved in a scanning mobility particle sizer (SMPS).

### (c) Optical size

The intensity of elastically scattered light from an aerosol particle is dependent on the particle size (refer to Chapter 12). For particles smaller than the wavelength of light the intensity is a monotonic function of particle size (Figure 10.8) [7]. An optical particle counter (OPC, Figure 10.9) records the intensity of light scattered from a single particle and estimates the particle size from a calibration curve obtained with monodisperse spherical particles of known size and refractive index. Uncertainty in refractive index may lead to errors of as much as 400%. Light scattering may be recorded at a fixed angle or integrated over a range of angles. The light source can be a monochromatic laser or an incandescent white light



**Figure 10.8** Variation in the scattering efficiency of a spherical particle at  $90^\circ$  with particle size and refractive index. The monotonic variation in scattering efficiency for small particles permits size determination from optical scattering intensity. If the refractive index is unknown or the particles are large, significant errors may result.



**Figure 10.9** An optical particle counter. Particles may first undergo condensational growth prior to counting.

source. The larger irradiance that can be obtained with a laser, as compared to a white light source, leads to minimum detectable sizes of 50 nm and 300 nm for these two light sources, respectively. Instruments capable of measuring particles in the size range 0.1–10  $\mu\text{m}$  are common, with size distributions recorded by determining concentrations in up to  $\sim 50$  size resolved ranges or channels. It is important that only single particles are illuminated and dilution of the aerosol may be necessary to achieve this. Coincidence errors may arise if the aerosol concentration rises above  $1000\text{ cm}^{-3}$ .

Particle heating can occur, leading to a decrease in particle size and increase in refractive index and inducing an error in the size determination. In addition, the particle shape, morphology and refractive index are typically unknown. With the calibration procedure usually performed with spherical particles, this can lead to uncertainties in the size determination. Information on particle shape and morphology can be gained by differential light scattering, examining the angular variation in scattered light intensity. Instruments now exist for measuring the angular scattering from a single particle, 1–10  $\mu\text{m}$  in size, for characterising particle shape.

OPCs can be combined with a CPC for examining particle sizes less than 10 nm. Particles <10 nm in diameter show a size dependent growth rate as a consequence of the influence of surface curvature on vapour pressure, referred to as the Kelvin effect. Thus, analysis of the scattered light intensity from particles leaving a CPC can provide size distribution information on particles smaller than 10 nm in diameter.

The integrating nephelometer measures the light scattered from an aerosol over a wide angular range, within 5–10° of both the back and forward scattering directions [5]. Simultaneous measurements of the total integrated intensity and the back scattering intensity can be made. Nephelometers are available that concurrently record the integrated scattered intensity at three wavelengths in the visible range, allowing information on the size distribution



to be attained. The preferential forward and backward scattering by coarse particles can lead to a significant underestimation of the coarse mode component in the size distribution. Furthermore, coarse particles may be rapidly lost through deposition within the instrument, and droplet heating may be significant, particularly for coarse particles with a large water component.

## 10.4 Determining Particle Composition

Aerosol composition can be determined from the analysis of aerosol samples collected on filters using a wide range of conventional techniques that identify and quantify chemical speciation and composition. Such measurements are susceptible to artefacts which may occur during sampling, transport or storage. Artefacts include the loss of volatile components from collected particles, and gas–particle and particle–particle reactions within a collected sample. Although real-time *in situ* analysis is preferred, there are very few techniques that can enable the direct on-line determination of aerosol composition.

### 10.4.1 Sampling and off-line analysis

Analysing aerosol particles collected on a filter or impactor substrate can allow a wide range of analytical techniques to be used. The time required for sample collection is dependent on the aerosol mass loading and sampling rates. Typically, sampling durations may extend beyond a day. In addition to the filter and impactor substrates discussed in Section 10.2.2, an aluminium foil substrate is used for sampling OC and EC: a carbon free substrate is essential as the total carbon loading is determined by measuring the amount of carbon dioxide released on combustion. If samples are to be analysed by ion chromatography, precleaned Teflon or Mylar substrates are used as filter or impactor substrates. Typical detection limits for anions are about  $0.01 \mu\text{g ml}^{-1}$  for anions and  $1 \mu\text{g ml}^{-1}$  for cations once the sample has been extracted by treatment with bidistilled water and ultrasonic treatment.

Teflon membrane filters are also used for non-destructive techniques such as X-ray fluorescence (XRF) analysis or proton induced X-ray emission (PIXE). Both analytical techniques are rapid, non-destructive and sensitive, providing a powerful approach for the elemental analysis of aerosol. In PIXE, a sample is bombarded with energetic ions, usually protons, and characteristic X-rays are emitted from the bombarded sample allowing elemental analysis. A major advantage of PIXE is the extremely high spatial resolution that can be achieved because of the highly focused nature of the proton beam, focused to a spot size of micrometres. Typically, up to 20 elements may be quantified from a single PIXE spectrum. Elements with atomic numbers greater than 15 can be analysed with PIXE with no difficulty, but elements with atomic numbers between 10 and 15 must be analysed with caution. Elemental concentrations as low as  $1 \text{ ng m}^{-3}$  can be analysed by PIXE. Care must be taken if the sample is thick, typically  $> 5 \mu\text{m}$ : the decrease in incident proton energy with thickness, and the attenuation of the X-ray signal and appearance of secondary X-ray fluorescence must all be accounted for.

Diffusion denuders are used for determining concentrations of semi-volatile compounds [5]. Sampled aerosol is passed through a conduit (cylindrical or annular) and gaseous

components are removed by diffusion to the conduit walls and subsequent reaction with a co-reactant selected to efficiently remove the gas phase component of interest. The aerosol particles flow through the tube unaffected and are collected on a filter. A further absorber is located downstream of the filter to collect any volatile components lost from the particles collected on the filter. This can enable an examination of the partitioning of volatile components between the gas and condensed phases. Devising denuder coatings that show similar efficiencies for removing a wide range of semi-volatile organic compounds is a major challenge. Deposition of particles  $> 1 \mu\text{m}$  and diffusional losses of particles  $< 100 \text{ nm}$  can lead to artefacts. It is possible to initially exclude all large particles with an impactor at the inlet, but this itself can perturb the gas phase composition.

Analytical techniques must be sensitive to examine the small amount of material that is collected. Other techniques that are used to provide an elemental analysis of aerosol particles include instrumental neutron activation analysis, atomic emission spectroscopy, atomic absorption spectroscopy, inductively coupled plasma emission spectroscopy and capillary zone electrophoresis. Techniques including high performance liquid chromatography (HPLC) and gas chromatography/mass spectrometry (GC/MS) are used to provide information on trace chemical constituents. Laser microprobe mass spectroscopy, laser induced breakdown spectroscopy and laser induced plasma spectroscopy can provide approaches for determining metals in aerosols. All three exploit the high irradiances achievable with high power pulsed lasers. The former can also provide information on speciation, detect trace OC compounds and provide some ability to discriminate between particle surface and bulk composition with a spatial resolution of  $0.5\text{--}3 \mu\text{m}$ . The ejected ion fragments following illumination with the laser pulse are analysed by mass spectrometry and detection limits are typically around a few ppm. Electron microscopy can provide valuable information on particle morphology and elemental composition. Scanning environmental electron microscopy can be carried out at pressures greater than 5 mbar, reducing some of the problems associated with losing volatile components. For any microscopy technique, collecting data from a statistically significant number of particles can be time intensive.

#### 10.4.2 Real-time analysis

Despite the advantages of on-line analysis in real time, avoiding sampling artefacts and the need for expensive laboratory procedures, there remain few techniques for determining aerosol composition directly *in situ* with high time resolution. Ion chromatographs can be configured to analyse on line the composition of aerosols using an automated particle collection system. This first involves the separation of gaseous and condensed phase species with a diffusion denuder.

A thermal/optical analyser for measuring OC and EC concentrations has been developed. Samples are collected on quartz-fibre filters, reflecting the affinity of quartz for organic constituents. OC compounds are volatilised by heating the filter to  $650^\circ\text{C}$  and the desorbed fraction is oxidised to  $\text{CO}_2$  with a  $\text{MnO}_2$  catalyst at  $1000^\circ\text{C}$ . This is then reduced to methane in a nickel firebrick and the methane is then detected in a flame ionisation detector (FID). EC can be detected in a similar manner, although it is oxidised at  $350^\circ\text{C}$  with  $\text{O}_2$ . Polycyclic aromatic hydrocarbons can be quantified using a laser induced fluorescence analyser or a fast-response flame ionisation detector. In the former technique, fluorescence intensity

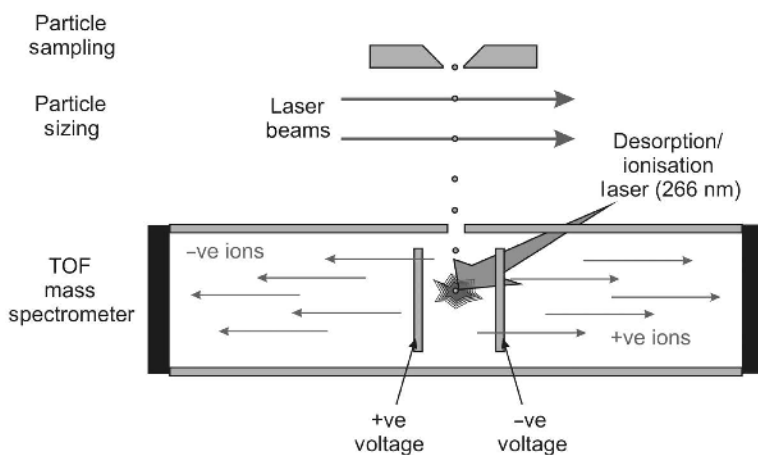
from a sample irradiated with a  $N_2$  laser can allow PAH concentrations to be determined with a sensitivity of typically  $<1 \mu\text{g m}^{-3}$ . EC in excess of  $0.5 \mu\text{g m}^{-3}$  concentrations can also be quantified with a photoacoustic soot sensor. Carbon black particles absorb energy from a laser beam and transfer heat to the surrounding gas phase. Modulation of the laser beam leads to modulation in the heat transfer and in the pressure pulses emanating from the heated particles. These can be detected with a microphone.

Sulfur species can be separated and quantified using a flame photometric detector, in which the  $S_2$  product produced in a hydrogen rich flame can be observed by fluorescence at 394 nm. A typical detection limit is  $1 \mu\text{g m}^{-3}$  with a response time of 1 min. Volatile and gas phase components are removed in a denuder and the remaining mass of the particle is probed. Speciation is possible by heating the aerosol to temperatures of 71, 142 and  $190^\circ\text{C}$  for the determination of sulfuric acid, ammonium sulfate salts and refractory sulfur species, respectively.

### 10.4.3 Single particle analysis

All off-line and on-line measurements described so far have a number of deficiencies [5]. Measurements are made on an ensemble of particles, probing the full or partially selected distributions of particle composition and size. It is impossible to determine simultaneously single particle size and composition. Measurements are often susceptible to artefacts that can occur during the sampling, transporting or analysis stage. Measurements cannot be performed with high time resolution.

Recent advances in mass spectrometry can now enable the simultaneous analysis of particle size, elemental composition and speciation in real time and *in situ* for single particles (Figure 10.10). Although a number of variants exist, they each consist of similar components. Particle size is determined first by an optical technique, either by integrated scattering intensity or by the time required to travel between two probe lasers. Laser desorption ionisation using a high energy pulsed laser is then used to desorb and ionise compounds present in



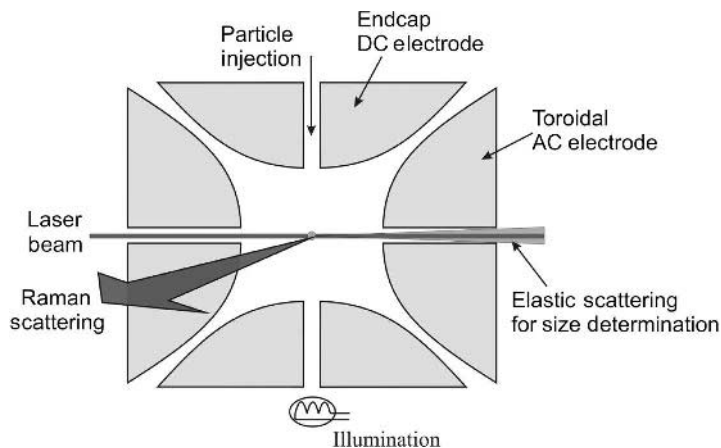
**Figure 10.10** A single particle time-of-flight mass spectrometer showing the sampling, sizing and mass spectrometer stages.

the particle. Detection is usually by time-of-flight (TOF) mass spectrometry. Dual polarity TOF spectrometers can be used to provide simultaneous information on the positive and negative ions formed in the ionisation step, yielding important information on chemical speciation and composition. The identification of key chemical tracers can enable the source of the particulate to be identified.

Although qualitative analysis by mass spectrometry is now routine, quantification can be difficult. This is particularly true for organic compounds, for which extensive fragmentation can occur at the laser desorption ionisation stage. Incomplete vaporisation of the aerosol particle can lead to a greater sensitivity for species on the surface of the particle than in the particle bulk. As an alternative, the desorption and ionisation can be achieved in two steps with separate lasers performing each process. Lower irradiances are required and this leads to less fragmentation of the desorbed compounds. Information can now be gained on the presence of elemental carbon, organic carbon, sulfates, nitrates, sea salt, dust, metals, etc.

In addition to mass spectrometric techniques, a number of approaches have been developed for sampling single particles non-intrusively and for probing particle size and composition using a range of optical techniques [8]. Solid or liquid particles can be levitated in an electrostatic trap or an acoustic trap, or in a focused laser beam. Electrostatic levitation was first employed by Millikan in the early 1900s to measure the charge of an individual electron and has become a powerful tool for trapping and characterising single aerosol particles. Contemporary instruments not only balance the gravitational force by varying the potential difference between top and bottom electrodes, but also confine the particle in the horizontal plane with an AC field applied to a toroidal electrode (Figure 10.11). Measurements of electric charge, mass and size can be routinely made.

Optical levitation can be achieved with a single focused laser beam [8]. The scattering force acts on the particle in the direction of propagation of the laser beam, balancing the gravitational force by radiation pressure. To achieve a stable trap, the laser power is controlled by an active feedback mechanism to balance any changes in particle size. More



**Figure 10.11** An electrodynamic balance. A single charged particle is trapped allowing single particle spectroscopy and dynamics to be studied. Composition can be determined from Raman scattering and size from elastic light scattering.

recently, single particles have been trapped in a three-dimensional single-beam gradient force optical trap, known as optical tweezers (refer to Chapter 13). A tightly focused laser beam not only leads to a scattering force but also a gradient force operating on a particle. Unlike the scattering force, the gradient force acts to push the particle towards the intensity maximum of the laser beam rather than away from it. If the light is focused to a tight waist, as is achieved with a microscope objective, the gradient force dominates the scattering force and a stable gradient trap is formed, confining the particle in three dimensions. It has been demonstrated that such a trap can be used to study particle size changes with nanometre accuracy. In addition, multiple particles can be simultaneously manipulated in parallel traps and particle coagulation can be studied.

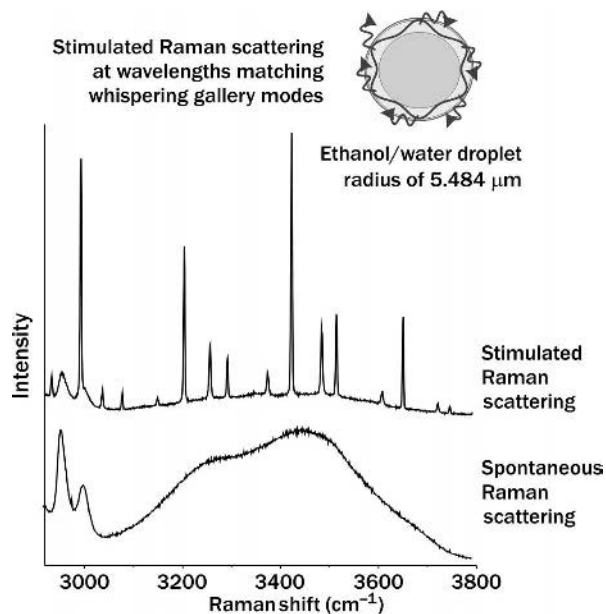
For particles trapped electrostatically, optically or acoustically, a wide range of spectroscopic techniques can be applied to characterise the particle. Raman, fluorescence and infrared spectroscopy are perhaps the mostly generally applied, providing optical signatures of droplet composition. Typical sensitivities for trace analysis *in situ* are  $10^{-3}$  M for Raman scattering from nitrate. Fluorescence detection is often used to discriminate between bacterial and other aerosols. Irradiation of single particles with a UV laser can allow real-time and *in situ* detection of fluorescence signatures from flavins, and amino acids such as tryptophan and tyrosine used as markers of bacterial material. Elastic light scattering can be used to determine the particle shape, refractive index and size, with nanometre accuracy for spherical particles.

The sphericity of droplets leads to the existence of morphology dependent resonances (also referred to as whispering gallery modes) [9]. These are formed at discrete wavelengths when the wavelength of the light is just right to form a standing wave within the droplet. Light undergoes total internal reflection within the droplet each time it encounters the interface and can circulate for timescales of nanoseconds leading to metres of pathlength in a particle that may be only a few microns in diameter. This can provide an optical feedback mechanism that can lead to stimulated Raman scattering or stimulated fluorescence at discrete wavelengths matching the resonant modes. A structured Raman or fluorescence band results that can allow both determination of droplet composition and also droplet size with nanometre accuracy (Figure 10.12).

The combination of this wide armoury of trapping and spectroscopic techniques can enable single particle dynamics to be interrogated over extended periods of time, with spatial and temporal resolution of composition and accurate determination of particle size and temperature.

## 10.5 Concluding Remarks

In this chapter we have illustrated the full breadth of information that can be gained in the analysis of aerosols, concentrating on the size, concentration and composition of the condensed phase. Over recent decades, the possibility of analysing single particles has become a reality enabling the microscopic factors that control the composition and reactivity of the aerosol to be studied with unprecedented detail. There remain many challenges in aerosol science and we have highlighted some of them in Section 10.1, particularly in the discipline of atmospheric science. In particular, the chemical and physical properties of multi-component and multi-phase inorganic/organic/aqueous aerosols remain poorly



**Figure 10.12** Comparison of the spontaneous and stimulated Raman scattering (SRS) fingerprints that result from an ethanol/water droplet. The SRS fingerprint can lead to the determination of both droplet size and composition with high accuracy.

characterised and understood. The development of new analytical techniques for field measurements and laboratory studies are set to play a crucial role in unravelling the chemistry that occurs in such a complex environment as the atmosphere.

## References

1. Spurny, K.R. (ed.) (1999) *Analytical Chemistry of Aerosols*. Lewis Publishers, Washington, DC.
2. Shaw, D.T. (ed.) (1978) *Recent Developments in Aerosol Science*. Wiley, New York.
3. Hinds, W.C. (1982) *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*. Wiley, New York.
4. Seinfeld, J.H. and Pandis, S.N. (1998) *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley, New York.
5. McMurry, P.H. (2000) *Atmos. Environ.* **34**, 1959.
6. Chow, J.C. (1995) *J. Air Waste Manage. Assoc.*, **45**, 320.
7. Bohren, C.F. and Huffman, D.R. (1983) *Absorption and Scattering of Light by Small Particles*. Wiley, New York.
8. Davis, E.J. (1997) *Aerosol Sci. Technol.*, **26**, 212.
9. Symes, R. Sayer, R.M. and Reid, J.P. (2004) *Phys. Chem. Chem. Phys.*, **6**, 474.

# Chapter 11

## Practical Rheology

*Roy Hughes*

### 11.1 Introduction

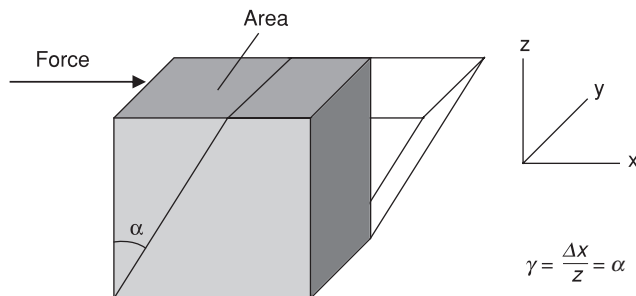
In this chapter we are going to adopt a hands-on approach to the study of rheology. This subject area is concerned with the science of the deformation and flow of matter. We will introduce the concepts with the minimum of mathematics and consider how to make measurements that are free from artefacts and relevant to many applications. We will investigate the relationship between the rheology of a material and its storage stability. Finally we will consider ways in which we may impart the flow behaviour we desire by the addition of colloid and polymer species to our formulation.

### 11.2 Making Measurements

It is easy to become daunted by the range of measurements [1–3] that can be performed by modern instruments. The unfamiliar terms and units that are used to enumerate the results can be intimidating to the first time user. The best way to overcome this difficulty is to grapple with some basic definitions. In the rest of this chapter we are only going to consider measurements in shearing flows. This type of flow is most commonly encountered in instrument laboratories. We should be aware that the processing of many materials results in complex flow patterns. These may involve stretching flows or forces normal to the flow direction and even turbulence. We should always bear this in mind when comparing instrumental data with an application for our material or formulation.

#### 11.2.1 Definitions

Imagine a small deformable cube of material. Suppose we apply a force to the upper plane of the cube whilst holding it steady. If the force is only gentle the cube will deform only slightly. The vertical edge will move through an angle  $\alpha$ . This is shown in Figure 11.1. There is a net stretch on the cube due to the applied force. For a homogeneous material the relative deformation is affine, which means it is essentially continuous throughout the body of the cube. The relationship between the force and the deformation is due to the physical properties of the cube and it is this that we are interested in. In principle we could measure



**Figure 11.1** The shear stress and shear strain on a cube.

the stretch of the cube and record this value for an applied force. This would be different for different sizes of cubes and is not a very convenient measure. The *relative deformation* of the cube is termed the (shear) *strain*,  $\gamma$ , and removes the dependence on the size of the body. It is the increase in the dimension in the  $x$  direction of the cube,  $\Delta x$ , relative to the height of the cube  $z$ . This equates for small deformations to the angular displacement. It is a dimensionless quantity. The force required to cause the deformation also depends on the area of the cube on which it acts. If we double the area we have to double the force to cause the same deformation. So we use (shear) *stress* as a convenient tool. This is the force divided by the area of the face of the cube. It has the unit Pa or  $\text{N m}^{-2}$  and is denoted by the symbol  $\sigma$ . The stress divided by the strain gives us a fundamental property of our material,  $G$ , the *shear modulus* as:

$$G = \frac{\sigma}{\gamma} \quad (11.1)$$

For a material which is described as *linear* the shear modulus maintains a constant value regardless of the stress or strain applied. Such a material is called a Hookean solid.

Whilst a simple elastic modulus might be an appropriate measure for rigid materials or perhaps some gels, many colloid, polymer and surfactant systems appear essentially fluid. It is difficult to practically arrange a simple cube to define our terms for these materials. We need to visualise a slightly more complex scenario. Imagine a fluid contained in a vessel with planar walls parallel to one another. We can consider just one small element of the vessel as shown in Figure 11.2.

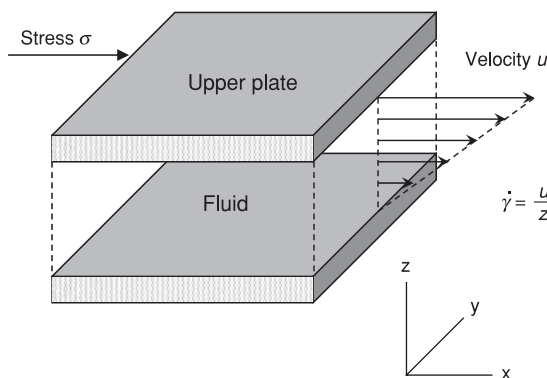
We will now apply a stress to the upper surface whilst maintaining the lower surface static. If the molecules forming the fluid are firmly attracted to the upper and lower plates, a velocity gradient will develop with the distance  $z$  across the gap. As the material is a liquid the displacement in the  $x$  direction is continuous. The strain is given by Equation (11.2) provided the gap is small and the velocity change is linear. The strain increases constantly with time  $t$ . At any time since the start of the experiment, the strain is given by the ratio of the distance moved  $x$  relative to  $z$ :

$$\gamma(t) = \frac{x}{z} \quad (11.2)$$

The velocity  $u$  is the rate of change of distance  $x$  with time. We can write this as

$$\dot{\gamma} = \frac{d\gamma(t)}{dt} = \frac{1}{z} \frac{dx}{dt} = \frac{u}{z} \quad \left( = \frac{du}{dz} \text{ for large gaps} \right) \quad (11.3)$$





**Figure 11.2** The shear stress and shear rate on two plates.

The term  $\dot{\gamma}$  is the shear rate or strain rate and for a constant applied velocity this is invariant with time. Thus the shear rate, which is constant with time, is a much more convenient term to use than strain, which is varying with time. The dot on the strain is called 'Newton's dot' and represents the time derivative of the strain. Shear rate has the units of reciprocal seconds,  $s^{-1}$ . In practice we set the shear rate by adjusting the conditions on the instrument we use. The most important relationship to emerge is that between stress and shear rate. The stress divided by the shear rate gives us a fundamental property of our material,  $\eta$ , the *shear viscosity* as

$$\eta = \frac{\sigma}{\dot{\gamma}} \quad (11.4)$$

For a material which is described as *linear* the shear viscosity maintains a constant value regardless of the stress or shear rate applied. Such a material is a Newtonian liquid.

### 11.2.2 Designing an experiment

Designing the correct experimental protocol is determined by what it is we would like to know about a sample, the sensitivity and the range of the method we are using. There are numerous measurement methods based on industrial standards. For example viscous flow can be monitored using a flow cup (ISO or Ford cup). Here a cup with a small hole formed in the base is filled with a liquid. The flow out of the cup through the hole is monitored with time. The stress acting on the fluid is due to gravity and the viscosity of the fluid determines the speed at which the liquid escapes. The shear rate is not single valued, but is determined by the complex flow patterns that develop as the fluid approaches and exits the nozzle. Some examples of cups are shown in Figure 11.3. Cups have found widespread and successful use in the paint industry for assessing coatings, although their use is not restricted to just this application. Drilling fluids are often characterised in terms of a pour point. The rate at which a viscous liquid in a beaker pours depends upon the gravitational stresses on the fluid and its viscosity. If we now reduce the temperature the viscosity of simple liquids increases. At some point it will no longer 'apparently' pour. We can define a pour point temperature



**Figure 11.3** Flow cups (pictures courtesy of Sheen Instruments Ltd).

for this material. Under these conditions it will prove hard to pump and filter. As a quality control tool and in difficult environments such methods have an important place.

These approaches become more difficult to use when we want to understand how a particular formulation performs under a range of flow conditions: particularly when we have no previous knowledge of the relationship between measurement and performance. This is partly because they impose complex flow patterns on the material and partly because they may not be operating on timescales appropriate to our needs. Viscosity and elasticity tend to be functions of stress and shear rate once a certain value is exceeded (i.e. non-linear). The aforementioned measures will do only a little to explore these phenomena and so tend to be more limited as tools for developing materials or understanding processes.

The best method to determine a rheological property for such a purpose is to control the force or deformation applied to the sample. Most of the colloidal materials we are interested in have a fluid or near fluid-like character. This means that the materials are not self-constraining and need to be introduced to a sample cell in order to hold their shape. Such sample cells have organically acquired the name *measuring geometries* or simply *geometries*. Figure 11.4 shows some designs of geometries. There are many variations. Typically they contain sample volumes from about 0.5 ml to 50 ml, or perhaps a little more. The choice of geometry can be critical and the wrong selection can result in meaningless data. Most viscometers or rheometers operate by using an electric motor to apply a known torque or known velocity. The stress, strain or shear rate that is applied to the sample depends upon the shape of the geometry attached to the motor and the performance of the motor.

Suppose we are investigating spraying or blade coating and we know that the sample will experience very high deformation rates. In order to explore how a new formulation may respond in the laboratory, we would want a high shear rate. Consider Equation (11.3); this shows that for a given velocity  $u$  (supplied by the motor), in order to obtain a high shear rate we need a small gap  $z$  (determined by the geometry). A typical high shear geometry is the Mooney Ewart which consists of a bob placed concentrically in a cup with a small gap between the two. One might achieve  $10,000 \text{ s}^{-1}$  with such a design. As the gap becomes



**Figure 11.4** A range of instrument geometries from Bohlin Instruments.

smaller the effect of any large contaminating particles becomes increasingly significant. Clearly particles can bridge the gap between the geometries, an extreme case, which would totally disrupt the measurements. There is a less extreme condition when particles have a diameter of about  $1/10^{\text{th}}$  of the gap size. The shear rate and the strain would not be affine and these effects alone begin to change the stress we would achieve in a larger gap at the same rate. It is equally more prone to particle jamming and local flow effects causing unusual and irreversible phenomena such as particle aggregation. This might be desirable if we are trying to mimic a process but it would not provide a well controlled measure of the rheological properties of a sample.

An overlooked problem with making a good measurement is that of wall slip. In order to get the shear rate we believe our motor and geometry selection has determined we need the assumption that the molecules would be firmly associated with the walls to be true (Section 11.2.1). If this is not the case, then we will form a *slip plane* between the geometry surface and the liquid next to it. Thus when we rotate the geometry it will move at a speed much faster than the neighbouring liquid and the shear again will be non-affine. The data will prove difficult to interpret. The extent to which a material might slip as a function of the velocity of the shearing surface is surprisingly poorly understood. Emulsions and foams are particularly prone to this problem as are high concentration particulate systems or skin forming materials such as doughs. The wetting characteristics of the material play a role. So if we roughen the shearing surface for example by sand blasting or attaching a double sided sticky tape, we can reduce the energy of contact between the fluid and the surface. This increases the 'wettability of the surface' and can eliminate slip plane occurrence. The local flow patterns can change, altering the hydrodynamics close to the surface. We can also change the materials of the geometries as a whole, switching between stainless steel and plastics or glass. This is also important for achieving a desired chemical resistance of our geometries. We should not always consider the adsorption of species as a positive benefit to our measurements. An example of this is the addition of high molecular weight polymers in order to reduce the drag between the walls of a pipe and the liquid it contains.

This reduces pumping energy and costs in long pipelines. The extent to which this effects polymer rheological measurements particularly in mixed systems in a laboratory instrument is not well understood and almost always ignored.

A related phenomenon to wall slip is fracture. We see examples of this in very viscous and slightly elastic slurries. Generally this occurs as the shear rate is increased during an experiment. The velocity reduces across the gap between the moving and the stationary surfaces. At some point after a given time and at some position across the gap a critical strain is achieved and the sample breaks creating a plane of slip between a rapidly moving and a slowly moving surface. The reason for the breakage depends upon the rheological properties of the material and so it is difficult to generalise when we would expect to see this.

Other considerations include the volume of the sample compared with the exposed area. Large areas and small volumes accelerate the effects of evaporation. This will of course lead to concentration changes but could also lead to films forming and bridging the moving and static surfaces, significantly increasing the apparent viscosity of the material.

High concentrations of polymer solutions can produce forces normal to the plane of rotation of the geometries. It is not uncommon to see materials ‘climb out’ of the geometry at moderate shear rates.

### 11.2.3 Geometries

Geometries are not always constructed to achieve a near perfect shear rate. For example in the DIN standard cup and bob (Figure 11.5), a well defined shear rate is achieved in the narrow gap between the walls. The conical base of the bob does not result in a constant shear rate across the geometry, thus leading to systematic errors in the data collected. A constant rate can be achieved using a narrow gap and a shallow cone angle. This is the basis of the



**Figure 11.5** A comparison of a cup and bob, and a cone and plate geometry.

cone and plate geometry. As the cone rotates the outer edge of the cone moves faster than the centre of the cone. Thus the velocity is not constant across the geometry. However by angling the cone such that the increase in speed corresponds to an increase in the gap, the shear rate can be made constant. The point of the cone must of course be truncated, this prevents a frictional drag between the cone and the plate. This is usually truncated by a distance of about 50–200  $\mu\text{m}$  and clearly this is an area where large particles could become trapped.

We need to account for another factor which can impinge on the quality of the rheological measurement. When we shear the fluid in a geometry, lines of flow develop. If we consider a small cubic element there will be a velocity gradient across the element. If we imagine the cube to be a rigid block, one side of the block would be travelling faster than the other. This results in a twisting motion or a vorticity. At low shear rates the viscous forces tend to dampen out this tendency but as the shear rate increases there is a tendency towards forming small vortices. These secondary flow patterns occur at a critical rate dependent upon the geometry, the density and the viscosity of the sample. The lower the viscosity, the greater the tendency to form these flows. The shear rate where we see these secondary flows can be estimated using the Taylor number. Once a critical value is exceeded, the shear stress at a particular rate is no longer a measure of viscous processes alone; it includes inertial properties.

When a very high shear rate is applied to the sample the vorticity becomes the dominant feature of the flow. At this point complex turbulent flow patterns develop and mixing occurs. The onset of this is determined by the ratio of the inertial to viscous forces. We term this ratio the Reynolds number and when we achieve high Reynolds numbers the flow is dominated by inertial rather than viscous processes [1, 3].

The lesson to learn here is that even the most wisely selected geometry can be used in regions where it produces data of very little value.

### 11.2.4 Viscometry

We can measure viscosity by applying a shear rate to a sample and recording the shear stress we obtain. Suppose for a moment that we have the perfect device so that it was able to apply an instantaneous velocity to the moving part of a geometry. A velocity gradient would develop with time across the sample and a *steady state* stress would be recorded. Simple liquids such as water or oils are Newtonian; that is the viscosity is independent of the shear rate. The molecules respond rapidly to the gradient, and the timescale over which the stress reaches a constant value is practically instantaneous on the timescales we are interested in. However, complex materials, the subject of our discussion here, are made of much larger species than single atoms or small molecules. As a result the timescale for the system to reach a steady state is comparatively long. This is an important observation since if we wished to record a steady state viscosity at *any* shear rate, we need to wait long enough for the larger species to respond.

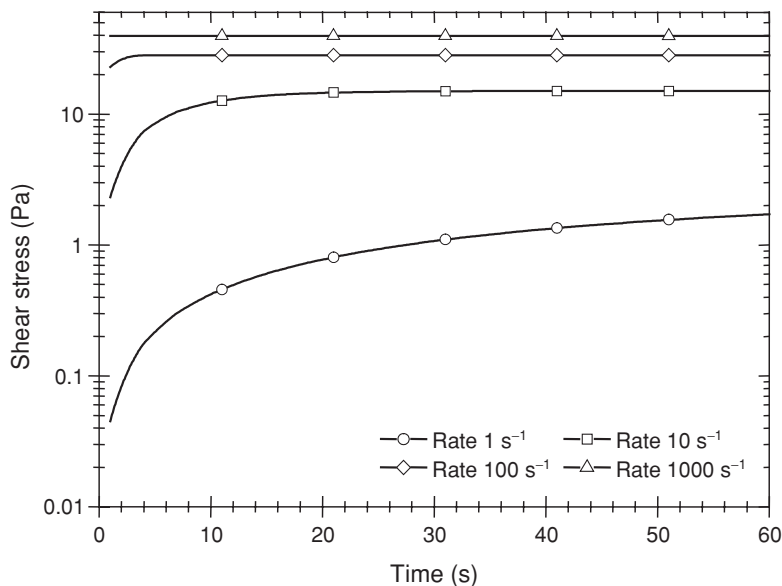
The rate at which steady state is achieved depends upon the diffusive properties of the species forming the system. The time required to diffuse a given distance is controlled by the diffusion coefficient in the sample. If we look at an experiment with a timescale much longer than some characteristic diffusion time we certainly expect our structure to have achieved steady state. We usually define a characteristic time,  $\tau$ , for a movement relative

to the dimensions of the species forming the system. So for example, for a dispersion of monodisperse particles with a radius  $a$ , we would write [3, 4]

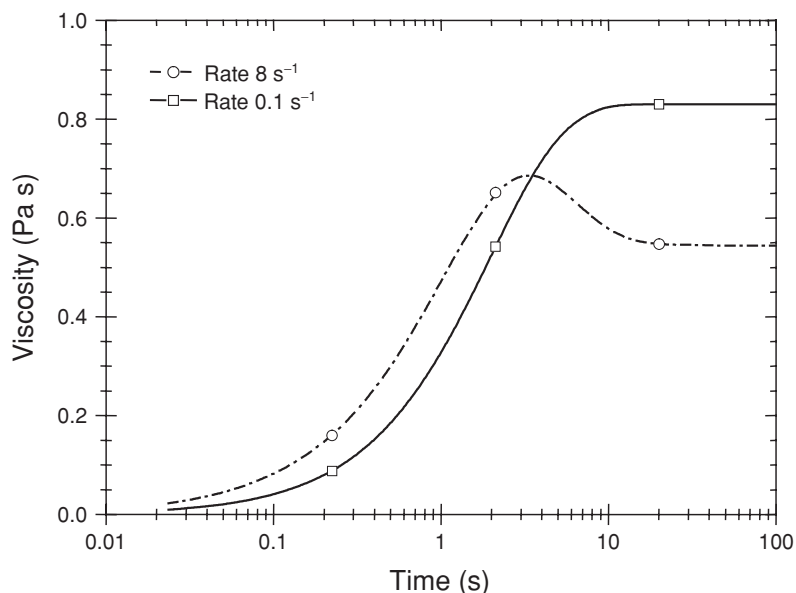
$$\tau = \frac{6\pi\eta(0)a^3}{k_B T} \quad (11.5)$$

where  $k_B T$  is the thermal energy of the particles and  $\eta(0)$  is the viscosity at low stresses or rates where the material is relatively unperturbed. As we suggested earlier the molecules forming the dispersion medium respond instantly. We could test Equation (11.5) and show that atoms are so small, of the order of Angstroms, that the characteristic time would be small compared to that of say a  $1\ \mu\text{m}$  silica particle. We have not made an arbitrary choice in selecting the particle length to characterise our system. Many concentrated particulate systems show a short range order akin to a true molecular liquid. The order typically occurs over a particle diameter or so. Thus when a particle moves a distance of the order of a particle dimension, substantial local order is lost. We are characterising the size and lifetime of the arrangement.

The practical outcome for this for our experimental design is that we must wait long enough for the structure to respond to our applied stress in order to get steady state data. Unfortunately Equation (11.5) is only a guide and also depends on our knowledge of the viscosity at a low stress or shear rate. We can illustrate this using some experimental data shown in Figure 11.6. This shows the stress that is achieved with time when four different fixed shear rates are applied to a dispersion. The first thing to notice is that as the shear rate is increased the shear stress increases. In order to see the data more clearly a logarithmic scale is used. At low shear rates the stress gradually increases to reach a steady state response, seen after about 1 min. As the rate is increased, a steady state is achieved more rapidly and at  $1000\ \text{s}^{-1}$  it has occurred faster than the instrument was able to readily record.



**Figure 11.6** A plot of shear stress versus time.



**Figure 11.7** An example of ‘stress overshoot’ illustrating difficulties in measuring time dependent rheology.

In order to achieve a steady state response it is important to allow enough time for the instrument to measure at each shear rate. It is possible to use feedback mechanisms in some instruments to sense when the viscosity reaches a plateau value. This can be convenient but should be used with caution. An example of why we should take care is shown by the response from a heavily entangled polymer solution [5, 6]. Figure 11.7 shows the stress response at a low shear rate (less than a few reciprocal seconds) and at high rates (a few reciprocal seconds). At low shear rates the viscosity, determined by the shear rate divided into the stress at any instant, increases to a plateau value. The polymer chains take time to rearrange to reach their steady state configuration. However when a high rate is applied the chains still attempt to follow the flow but they are unable to change their mutual entanglements fast enough to respond to the higher flow rate. They begin to store energy in highly distorted chains and much more stress is required to achieve a similar rate. This leads to a peak in the viscosity or a stress overshoot. Eventually enough shear energy is input to the system to enforce a new steady state conformation and the viscosity falls to a plateau value. A material which loses its viscosity with *time* when a single shear rate is applied and which recovers the viscosity once the shearing is stopped is termed *thixotropic* (the reverse being *anti-thixotropic*). We can see from our polymer example (Figure 11.7) that this simple classification is not always adequate. It is important to clearly distinguish the difference between time dependent responses at any one shear rate separate from the steady state behaviour we observe at different shear rates. It is very common for *shear thinning* and *thixotropy* to be confused. This can lead to very serious formulation design problems.

Although steady state behaviour is often the most desirable form of characterisation, sometimes the time required to achieve this state can be prohibitively long. It is quite feasible for one hour or more to pass before an equilibrium response is achieved at low shear

rates. Under such circumstances a *creep* (Section 11.3.3) test is much more appropriate for obtaining information.

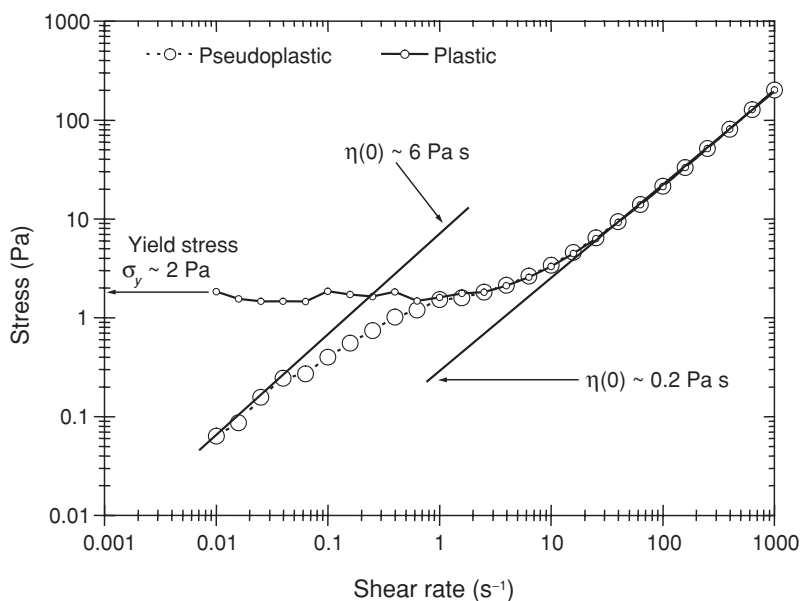
It is worth considering that there are some circumstances where we may not wish to achieve a steady state response. For example we may be imitating a process where our sample is transported from a trough to a blade coater and the fluid is rapidly deformed. A steady state response is probably not a relevant measure here.

One experimental method, which is commonly available on viscometers, is a sweep of shear rates. These experiments are designed to increase the shear rate from say  $0 \text{ s}^{-1}$  to  $500 \text{ s}^{-1}$  and then back to  $0 \text{ s}^{-1}$  in a known time regardless of the steady state response. This can result in very complex looking curves, optimistically called thixotropic loops. Although the experiment can be performed in a very controlled manner the data are very difficult to interpret. Not only is it unclear at which rates, if any, a steady state response is achieved but also the material has been exposed to a complicated *shear history*.

The rheological properties depend on the shear history of the sample. This means that the response depends upon all the previous deformations and stresses applied to the sample. This also includes loading the sample into the geometry. Often the best we can do is to avoid applying high shear on filling the geometry and leaving the sample a reasonable time to recover. However what constitutes a 'reasonable time' is a matter of opinion and patience.

### 11.2.5 Shear thinning and thickening behaviour

By using a range of geometries it is possible to cover a very wide range of shear rates and shear stresses. Accordingly the best way of representing this data is to use logarithmic scales. Typical example plots for two different silica dispersions are shown in Figure 11.8. They



**Figure 11.8** A double logarithmic plot of shear stress against shear rate.



typify two different classes of behaviour. For both samples we see that as the shear rate is reduced the stress reduces. One curve goes through a point of inflection and then continues reducing. At both high and low shear rates the data go to a slope of unity. This means that the shear stress divided by the shear rate gives a constant viscosity both at low and high shear rates. The material will flow at both low and high shear. This defines a low shear rate viscosity  $\eta(0)$  and a high shear rate viscosity  $\eta(\infty)$ . This type of flow curve characterises a *pseudoplastic* material. The other sample however does not display a low shear viscosity but a yield stress  $\sigma_y$ . If a stress is applied which is less than this value no flow occurs. This is a *plastic* material.

We can establish the viscosity on a log–log plot by extrapolating to the value on the stress axis when the line fitted to the shear stress–shear rate curve intersects with a shear rate of  $1 \text{ s}^{-1}$ . One feature of these two sets of data is the similarity of the high shear viscosity. If the data are re-plotted on a linear–linear scale the two samples appear almost identical and it would be hard to distinguish the material with a yield stress. A yield stress essentially implies the material is solid-like at low rates and stresses. The presence of a yield stress is important for storage stability and this is considered further in this chapter.

A fuller characterisation of these curves is usually achieved by fitting the data to a mathematical model. This allows us to compare trends in the data for materials of varying composition using a limited number of constants. Fitting flow curves to a model and extracting the constants as functions of composition also allow systems to be compared more readily than plotting a large number of curves side by side. These models describe the phenomena and are termed, unimaginatively, phenomenological models [3].

Plastic models:

$$\sigma = \sigma_{by} + \eta_{pl}\dot{\gamma}, \quad \text{Bingham} \quad (11.6)$$

$$\sigma = (\sqrt{\sigma_c} + \sqrt{\eta_{pl}\dot{\gamma}})^2, \quad \text{Casson} \quad (11.7)$$

$$\sigma = \sigma_{hb} + (\eta_{pl}\dot{\gamma})^{-n+1}, \quad \text{Herchel–Bulkley} \quad (11.8)$$

The yield stress is given by  $\sigma_{by,c,hb}$  and  $\eta_{pl}$  is the plastic (high shear) viscosity.

Pseudoplastic models:

$$\eta(\dot{\gamma}) = \eta(\infty) + \frac{\eta(0) - \eta(\infty)}{1 + (\beta\dot{\gamma})^n}, \quad \text{Cross}, \quad (11.9)$$

$$\eta(\sigma) = \eta(\infty) + \frac{\eta(0) - \eta(\infty)}{1 + (\alpha)^m}, \quad \text{Kreiger}, \quad (11.10)$$

where  $\alpha$ ,  $\beta$ ,  $m$  and  $n$  are constants. It is not always possible to measure either the high or low shear rate behaviour. Data can still be fitted using power law behaviour:

$$\sigma = A\dot{\gamma}^{-n+1}, \quad \text{Ostwald–de Waele} \quad (11.11)$$

At very high shear rates or stresses with viscous slurries it is possible to observe a log jamming effect and this can lead to increases in viscosity. Trying to establish the steady state nature of these is extremely difficult.

## 11.3 Rheometry and Viscoelasticity

In Section 11.2.2 we suggested that most of the materials of interest to us in this discussion had near fluid-like character. This intentionally vague terminology allowed us to dismiss concerns about the physical state of a material. The ‘true state’ of a material, whether it is solid or liquid-like can be of great importance to us. We will now start to explore this in more detail in the next sections and consider its relationship to sedimentation, one form of storage instability.

### 11.3.1 Viscoelasticity and Deborah number

Suppose a strain is imposed on a solid sample by the application of a stress. The size of the stress is determined by the shear modulus (Equation (11.1)). Now just imagine there is a little fluid character to our solid and the molecules that form the material begin to rearrange and move relative to each other. The energy stored in the structure is dissipated by the flow and the stress begins to reduce. The rate at which the sample relaxes is determined by the balance of elastic and viscous processes. We can define a relaxation time,  $\tau$ , as the ratio of the viscosity to the elasticity:

$$\tau = \frac{\eta}{G} \quad (11.12)$$

This time is related to that defined in Equation (11.5) which measures a diffusion timescale for particles to begin to lose their local order. If we were to make experimental observations at times much shorter than this, we would see a more elastic character in our sample, and at times much longer, we would see a more viscous character. We can define a ratio of the relaxation time to the experimental observation time,  $t$ :

$$De = \frac{\tau}{t} \quad (11.13)$$

This ratio is the Deborah number  $De$  and this is a measure of the tendency of a material to appear either viscous or elastic. We can classify the material into three classes of behaviour:

$$\begin{array}{ccc} De \gg 1, & De \sim 1 & \text{and} & De \ll 1 \\ \text{solid-like} & \text{viscoelastic} & & \text{liquid-like} \end{array}$$

So when the observational time approaches the relaxation time of the material it will display both elastic and viscous characteristics and is termed *viscoelastic*.

### 11.3.2 Oscillation and linearity

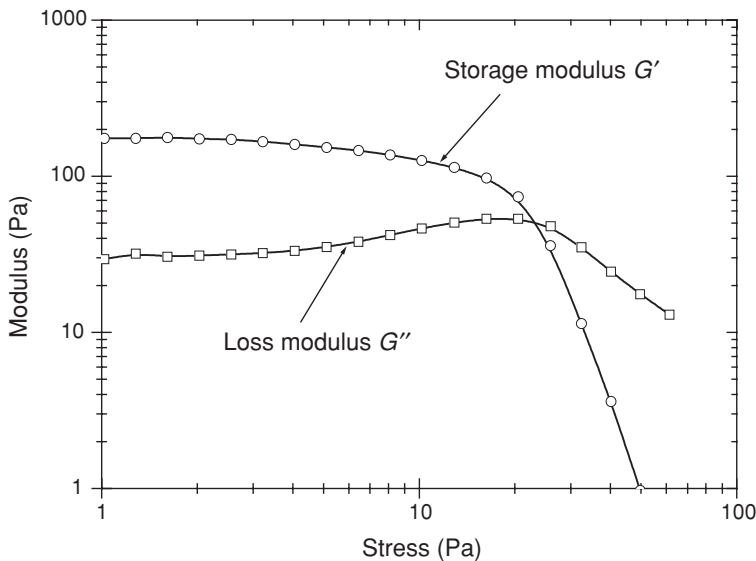
Characterising viscoelastic materials requires very subtle tests [5, 7, 8]. It is most convenient to work in a linear regime. This means that the shear modulus and the shear viscosity are independent of the strain or the stress used to measure them. They are only functions of time or frequency. There is a very good reason for doing this, which is that when we apply small deformations or forces we do not disrupt the structure of the sample too greatly. This makes it easier to interpret the data.

Normally the first experiment we would perform is one to test the limits of linearity [3]. We would probably need to replace the sample in our geometry after this test to ensure that the structure is not permanently damaged by the experiment. A typical test is a sweep of stresses applied to a sample at a fixed frequency. The material responds by developing an oscillating strain. If the material were purely elastic the strain and stress would be coincident. This means a peak in the oscillating stress is accompanied by a peak in the oscillating strain. The waves are said to be in phase. Now as we introduce some viscous character to the sample some energy is dissipated and the stress and strain waves begin to mismatch. A phase difference has occurred between the waveforms. We can utilise the phase difference to divide the waveforms into two terms:

- A strain in phase with the stress, an elastic energy stored component
- A strain  $90^\circ$  out of phase with the stress, a viscous energy dissipation component.

The ratio of the peak in the stress to the peak in the strain is the storage modulus  $G'$  for the in-phase component and the loss modulus  $G''$  for the out-of-phase component.

The protocol used for the measurements consists of selecting a frequency, say 1 Hz and applying this to the sample at a given stress. The procedure is then repeated at a higher stress and so on. The storage and loss moduli should be constant (or in a practical sense nearly so) until a critical strain or stress is reached. This sort of behaviour is shown in Figure 11.9 for a concentrated emulsion in the form of a cream. Here we clearly observe the change in the storage modulus as the critical stress is exceeded. Ideally we would like to work at a stress which is in the linear region so in the example (Figure 11.9) a value below 5–10 Pa would be a good choice. The question of how low a stress we should use is a subtle balance. If we pick a value too close to the end of the linear region, this can lead to problems since the application of a stress in an apparently 'linear' region over a very long period of time



**Figure 11.9** A shear stress sweep for a cream.

can occasionally result in slow structural breakdown. However if we select a stress which is too small the strain can reduce to a point where the displacement is small and the signal contains a high level of 'noise'.

Suppose we take the critical stress and divide this by the storage modulus; at the point where the breakdown in linearity occurs we get a strain of 0.05 or 5% (i.e.  $G' = 200$  Pa and  $\sigma = 10$  Pa; this is a typical value for a dispersion or an emulsion. If the system is very aggregated we might expect a lower strain will lead to a loss in linearity. Many polymers can sustain much higher strains.

We have not considered the choice of the geometry for viscoelastic experiments in any detail. It is important that we should consider all the aspects mentioned in the previous section. There is an additional concern which is the issue of geometry inertia. Geometries have a finite mass and so they require energy to enforce their motion. We can best aid the efforts of our instrument manufacturer (in their designs to allow for the energy dissipated in the motion of the geometry) by selecting lightweight geometries. If we want low stresses we need a large sensing/application area and in order to reduce the weight of the geometry we can use less dense but rigid materials such as titanium. This will improve the quality of the low stress and strain data.

### 11.3.3 Creep compliance

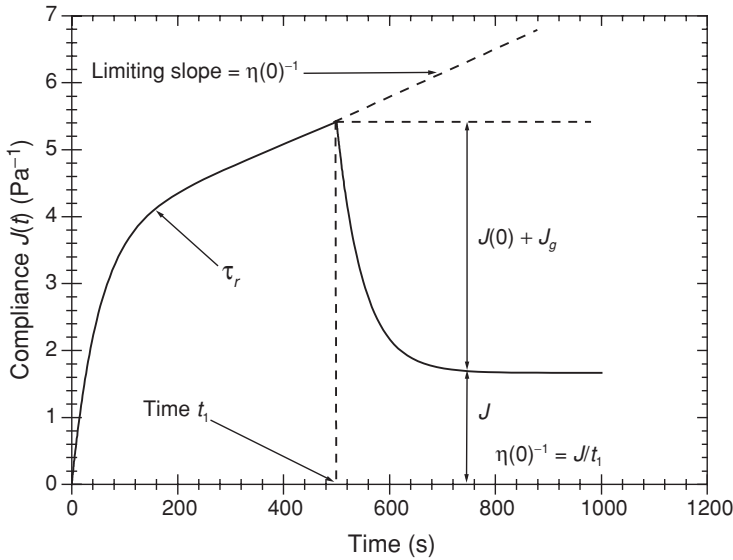
Sensorial testing tends to show that we have a good intuitive feel for the viscosity of a material. We can compare two materials and providing there is a reasonable difference between their viscosity and they are not too thick or thin we can say which is more viscous by feeling them. We do not have such a good 'feel' for elasticity, the other key property of a viscoelastic material. For example, if we compare a block of metal with a bowl of jelly we will find there is a tendency to say the jelly 'is more springy', 'it is more elastic'. This common usage of the word is entirely at odds with the scientific usage of the word. Whereas a gel may have a shear modulus of 100 Pa say, the metal block has a shear modulus of 100 GPa (1,000,000,000 times larger). In fact it is better if we think about the quantity compliance. The shear compliance is denoted by the symbol  $J$ . It is the ratio of the shear strain to shear stress:

$$J = \frac{\gamma}{\sigma} \quad (11.14)$$

It is reciprocally related to the shear modulus. In a creep compliance experiment we apply a stress to the sample and follow the compliance with time,  $J(t)$ , after the application of a shear stress. This experiment is a very good method for establishing if a material is essentially viscous or elastic.

### 11.3.4 Liquid and solid behaviour

The application of a step stress to a constrained material results in a displacement of part of the geometry. Initially an instantaneous compliance,  $J(\infty)$ , is recorded and then the strain grows to show a change in compliance with time. In practice the instantaneous response can be difficult to observe: in an experiment it depends on how the signal is filtered and how



**Figure 11.10** The creep compliance curve of a viscoelastic fluid.

quickly the strain can be applied by the instrument. This test is referred to as a creep test [5] and a typical curve for a fluid is shown in Figure 11.10. The remainder of the creep curve for a viscoelastic material is made of two types of processes. We obtain a curve as the body is deformed because the components which form the material rearrange and take time to do so. The response is not instant – it is slowed and is normally termed a retarded response. It has a characteristic retardation time  $\tau_r$  and an elastic component or compliance  $J_g$ . At longer times we generally see one of two processes occurring.

For a material which is pseudoplastic and has a low shear viscosity, the strain and the compliance will keep increasing all the time the strain is applied. This is shown in Figure 11.10. At long times the strain increases in direct proportion to the time. The material has attained a constant shear rate across it; the strain divided by the time is the shear rate:

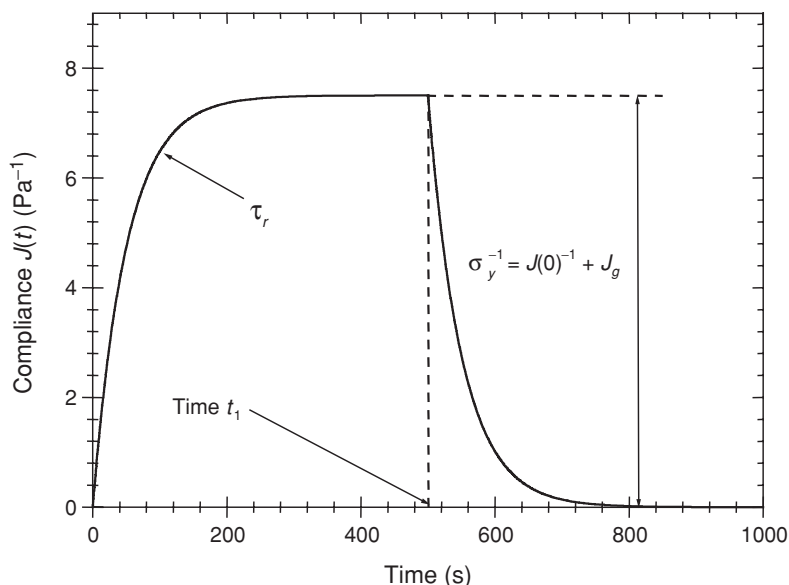
$$\dot{\gamma}_{t \rightarrow \infty} = \frac{\gamma}{t} \quad (11.15)$$

The viscosity is the shear stress divided by the shear rate. We may express this in terms of the strain,

$$\eta(0) = \frac{\sigma}{\dot{\gamma}_{t \rightarrow \infty}} = \frac{\sigma t_{t \rightarrow \infty}}{\gamma} \quad \text{or} \quad \eta(0) = \frac{t}{J(t)_{t \rightarrow \infty}} \quad (11.16)$$

As shown above, we can replace the strain divided by the stress with the compliance. This equation shows us that the gradient of the compliance versus time curve gives us the reciprocal of the viscosity. The point being made here is that if we run a creep experiment for an appropriate time and determine the slope, we can determine the viscosity.

Suppose we now remove the stress after a time  $t_1$ . If all the rearrangements of the species forming the material have occurred by the time the stress is removed, the geometry will stop moving and the strain will be constant. If the time  $t_1$  is too short compared to the retardation



**Figure 11.11** The creep compliance curve of a viscoelastic solid.

time  $\tau_r$  not all the rearrangements of species will have occurred and the material will still possess some elastic properties. When the stress is removed the material will recover some of its original shape. A portion of the original displacement will be returned and the compliance reduced. This portion of the curve is known as creep recovery.

If the material is a viscoelastic liquid we can *never recover all* the strain, some will be lost through the process of flow. If it is a viscoelastic solid it will not possess a zero shear viscosity but will have a yield point. This material will recover all the strain applied to the sample provided the yield stress is not exceeded. Such a response is shown in Figure 11.11.

For very viscous materials it can be difficult to distinguish between the two states. Mathematical analysis of the curve can help the distinction to be made. The plots also show several ways of obtaining viscous and elastic responses. Creep tests do have some direct analogues in terms of applications. For example the levelling of a film under the action of a gravitational stress has some similarities to the response of a material under a step shearing stress. Anomalous behaviour in one is reflected in the other, so it can prove a good tool. This rule tends to work less well with high loadings of polymeric materials. Another use of this technique is suggested in the following section.

### 11.3.5 Sedimentation and storage stability

Rheology is a very important discipline when it comes to distinguishing between the states of matter. As we have seen in the previously described creep test, we can differentiate between viscoelastic materials that have viscous or elastic properties at low shear stresses. This is a very useful indicator of the ability of a material to support solids [3]. Imagine we have a

dispersion of particles. Let us suppose that the particles are denser than the medium. The gravitational force  $g$  will cause particles of radius  $a$  to sediment. An isolated particle will sediment with a velocity  $v$  proportional to its frictional drag in the medium and the density difference  $\Delta\rho$  between the particle and the medium:

$$v = \frac{2}{9} \frac{\Delta\rho g a^2}{\eta_o} \quad (11.17)$$

The viscosity in this expression is the solvent viscosity  $\eta_o$ . This equation applies equally to creaming where the particles are less dense than the medium and migrate to the surface of the fluid.

If, in a sedimenting system, we were to increase the concentration of the particles, they will interact with each other slowing their velocity of descent. In a pseudoplastic material with a zero shear rate viscosity we can assume the gravitational stresses are too small to induce thinning, this gives us an expression which has a zero shear rate viscosity as opposed to the solvent viscosity:

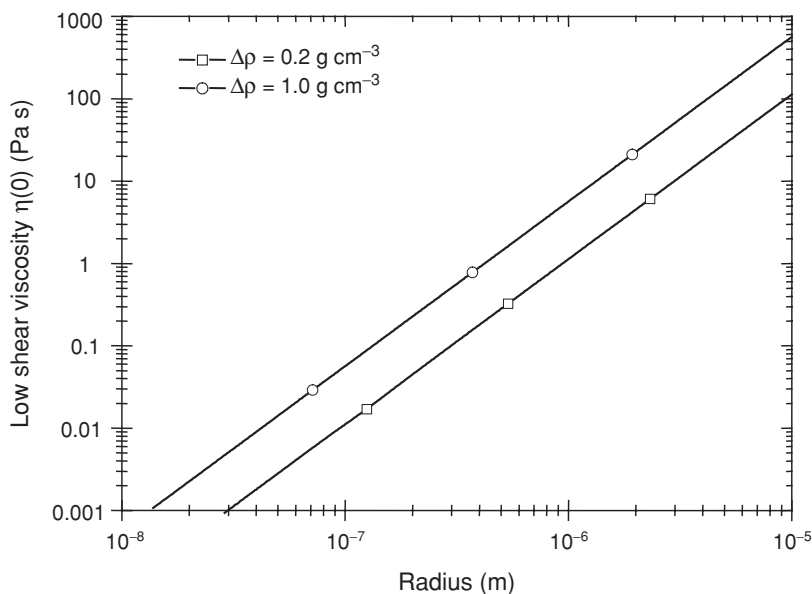
$$v = \frac{2}{9} \frac{\Delta\rho g a^2}{\eta(0)} \quad (11.18)$$

A cursory inspection of the equation shows that the higher the low shear viscosity the more slowly the material will separate. Therefore storage stability will tend to increase with higher low shear viscosity. We can do a little better with our prediction. Let us set ourselves a storage stability rejection criterion. We will reject our product if it sediments more than 1 mm in 1 month. This gives us a velocity of nearly  $4 \times 10^{-10} \text{ m s}^{-1}$ . We may now use Equation (11.18) to determine the viscosity we require to achieve this stability. Figure 11.12 shows a plot of particle radius on the  $x$  axis versus the viscosity on the  $y$  axis required to achieve our desired storage stability. This plot shows that small particles, nanoparticles, require a much lower viscosity than larger species. For a particle  $2 \mu\text{m}$  in diameter and with a density difference between the particle and the medium of only  $0.2 \text{ g cm}^{-3}$ , the viscosity required to stabilise the system is already at  $1 \text{ Pa s}$ : 1000 times that of water. For some processes this is unacceptably large. The solution to this difficulty is to reformulate the system to introduce a yield stress.

A plastic material has a yield stress. This has probably arisen through the formation of a network of interconnecting species. Suppose we know that gravity tends to cause the sedimentation of particles in the system. We may calculate the stress on these particles due to the gravitational forces. This is controlled by the particle mass, their density difference compared to the medium and the area over which the gravitational force is acting. Considering all these elements we obtain

$$\sigma_y > \frac{\Delta\rho g a}{3} \quad (11.19)$$

So if the sample has a yield stress greater than the value calculated on the right-hand side of the expression, we have good reason to expect some stability. We should not treat the results from this calculation as providing us with an exact solution, but more as a very good guide to the yield stress we require in order to gain storage stability.



**Figure 11.12** A plot of the low shear viscosity required to achieve stability in 1 month as a function of particle size and density.

There is an important practical limitation to this approach. Let us consider a typical limit of sensitivity for an instrument as 0.01 Pa. This means we cannot determine yield stresses below this value. Now take a  $2\ \mu\text{m}$  diameter particle in water with a density of  $1.1\ \text{g cm}^{-3}$ . The yield stress to keep this stable should exceed 0.0003 Pa: 33 times smaller than we can measure. Even if the density difference was 10 times larger it is still 3.3 times smaller than we can measure. One lesson to be learnt here is that it can be difficult to be truly sure of the stability of our system from a single measurement. Often it is the balance of evidence from viscoelastic studies at low stresses that provides the clues as to whether the material will store well. Another lesson to be learnt here is the inadequacy of flow curves when it comes to establishing yield stresses for sedimentation control. Attempting to extrapolate from a flow curve, where stresses of say 100 Pa or 1000 Pa have been used, down to a stress less than 0.1 Pa is fraught with errors.

Complex pastes can often give rise to another form of 'phase separation' called syneresis. This can be mistaken for sedimentation. This is understandable as it gives rise to a separated fluid on the upper surface of the paste. We see this sometimes in English mustard. In a container it can have the appearance of sedimentation. The process that is normally responsible for this is a network rearrangement between species in the paste. Following processing, the material begins to relax and rearrange to give a new network structure. The new network occupies a smaller volume than the original mixture, squeezing out liquid as it forms. It can differ from sedimentation visually in that we can sometimes observe the paste dewetting from the container to produce a lubricating layer between the vessel and the solid body of the paste.



Opaque systems present another difficulty. Larger particles separate more readily than smaller ones. So size polydispersity can lead to an uneven distribution of particles that is difficult to observe by eye since the opacity may appear unaltered throughout the system. This can lead to thick gooeey sediments developing undetected. One should consider Equations (11.18) and (11.19) for all the species in our system.

Finally, we should also be aware that as the system sediments, the concentration of species increases down the vessel. The rheological properties tend to increase in magnitude. This can halt sedimentation, so we can observe over a few days that a clear layer develops, gets a little larger and then becomes static. The system reaches equilibrium. An old trick is to sample this new concentration in the sedimented layer and reproduce it: in most cases our new rheology should prove to be just right to maintain stability.

## 11.4 Examples of Soft Materials

Many of the systems of practical interest to us consist of complex mixtures of materials. They may carry an active ingredient and require the correct rheological properties to deliver it effectively. Alternatively the rheological properties may be an intrinsic part of their application such as in paint. It would be convenient if these complex mixtures of materials can be blended to give ‘exactly the right’ properties. So suppose we have a blend of a polyacrylate dispersion and a surfactant solution. If we know the viscosity of the dispersion is 100 mPa s and the viscosity of the surfactant solution was 3 mPa s it would be very convenient if we found the combination was 103 mPa s or 51.5 mPa s or obeyed some simple combining rule. Unfortunately such relationships rarely hold and it is very difficult to predict the viscosity of many blends of materials. The main reason for this is straightforward. Colloidal materials tend to show interfacial activity and when we introduce one material to another new relationships develop between the components. It is no longer a simple mixture of the original components. This leads to complex changes in the rheological fingerprint of the material. The key thing is to understand the interactions between the components that form the material, then we can begin to design a formulation to give what we want.

We can ask ourselves a question: what are the main features that control the rheology of a system? Clearly it depends in detail on whether we are discussing polymers, particles or surfactants and so forth. We can identify some key embracing rules that determine the magnitude of the properties. These are as follows [3–6]:

- the number of interactions between species per unit volume (the number concentration, for example entities per ml)
- the size of the entities
- the strength of the interactions (for example in a mixed polymer surfactant system if the polymer is attracted to the surfactant and how strongly)
- the spatial arrangement of the species.

It is quite a tall order to have all this information in sufficient detail even in highly contrived model systems. In general the way we should approach designing or controlling the rheology is to have a good grasp of what we want and the type of colloidal interactions that can achieve it.

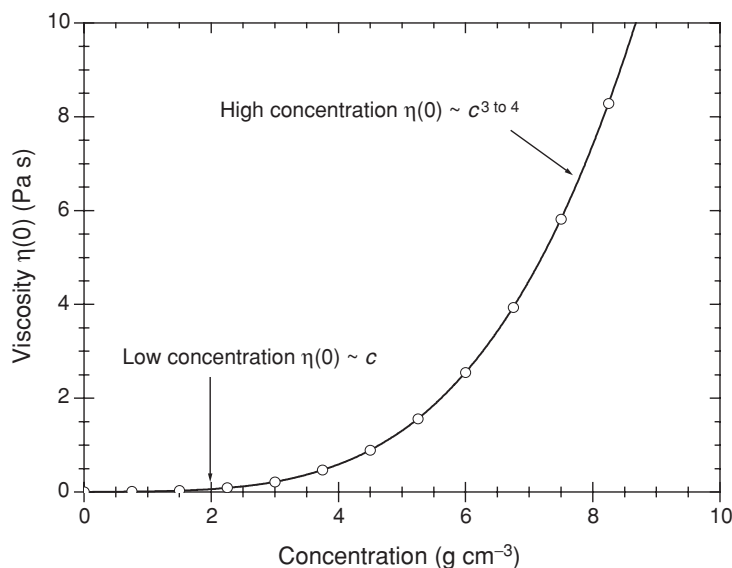
### 11.4.1 Simple particles and polymers

Frequently the rheological properties of formulations are adjusted by the addition of thickeners. These come in two forms: particles and polymers. Let us begin with a simple model for the flow of these species. Imagine a system of species that only interact through the flow lines they generate as they move through the medium and the rearrangements they can undergo. These are referred to as non-interacting species since the interactions are not influenced significantly by other types of interaction forces. Particles of this nature are referred to as hard spheres, and polymers are described as ideal.

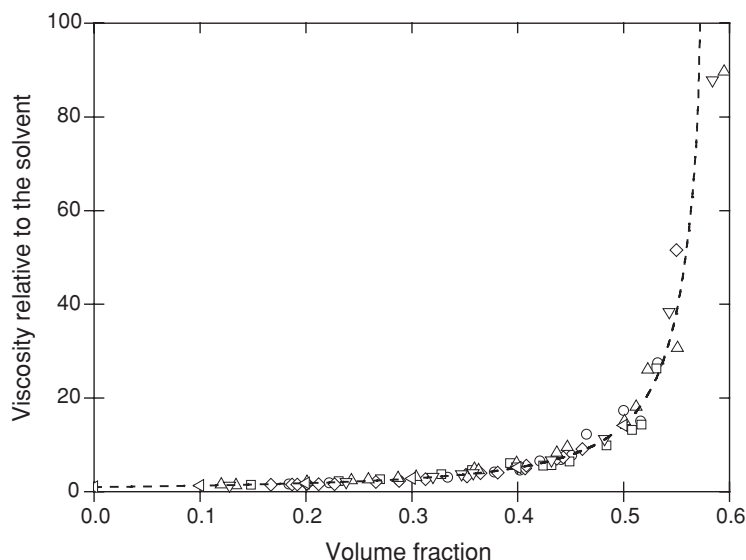
Polymer chains will change their conformation both internally and with respect to each other. The primary cause for the flow we observe in non-functionalised polymers is the entanglement of the chains [5, 6]. The more concentrated the system, the more entangled the chains become. There is a critical molecular weight, hence size, where this happens. The viscosity changes gradually with concentration, roughly linearly. At high concentrations power laws of 3–4 are observed before the systems become too viscous to measure using most practical methods. The onset of the higher power law behaviour is dictated by a critical concentration where polymer entanglements dominate the arrangement of chains. The trend in viscosity is shown in Figure 11.13.

Particles are restricted to rearranging their spatial distribution since they are too rigid to allow internal rearrangements to occur [3, 4]. The concentration of the species determines the viscosity at any point. The concentration dependence of the viscosity for particles shows a similar profile to that of polymers. We usually use the volume fraction to represent our concentration profile. A typical curve is shown in Figure 11.14.

The viscosity increases more rapidly with concentration for dispersions than it does for polymers. The rapid increase occurs near the maximum packing fraction of the dispersion.



**Figure 11.13** The low shear viscosity of a polymer as a function of concentration.



**Figure 11.14** The viscosity of a dispersion at high shear rates as a function of volume fraction.

This is the point where the particle network becomes space filling; it occurs at a volume fraction called the maximum packing fraction, typically in the range  $\phi_m = 0.5\text{--}0.7$ . This applies to both the high and low shear rate limiting viscosity. The packing is always smaller at the low shear rate limit than at the high shear rate limit.

In general the shear thinning profile of the systems has a similar form as illustrated in Figures 11.15 and 11.16.

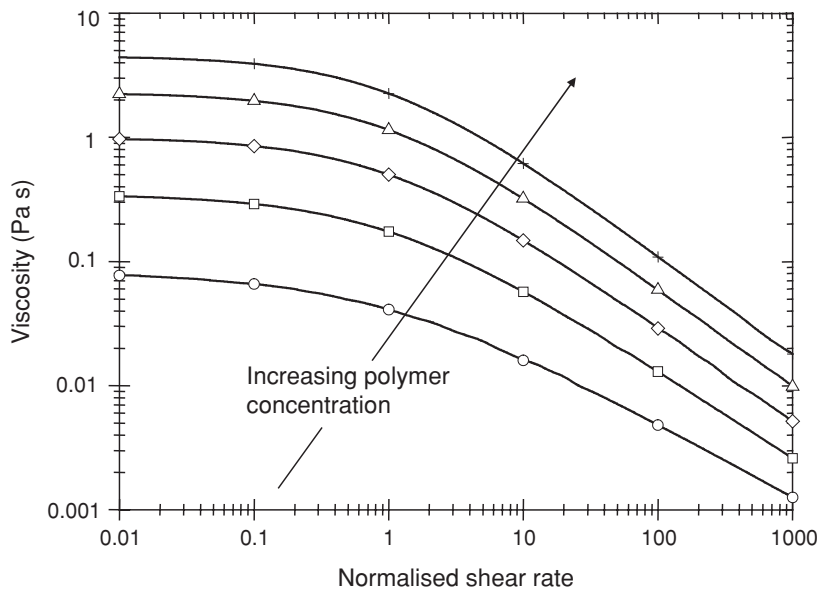
The key points illustrated by these flow profiles are as follows:

- It is easier to achieve a high shear rate limiting viscosity with particles than with polymers.
- The rate of change of viscosity with shear rate is more rapid with particles than with polymeric species.
- The low shear rate viscosity and the degree of shear thinning are linked in solution polymers; as the low shear viscosity is increased the degree of shear thinning increases.
- Significant shear thinning is only achieved with high concentrations of particles and often close to a point where small changes in concentration lead to massive changes in viscosity.

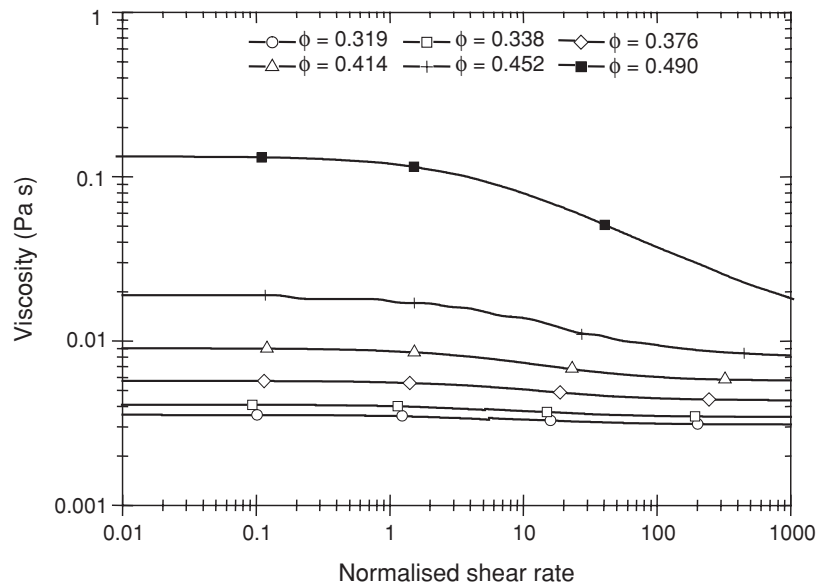
The latter condition is not usually a good position to find yourself in with a formulated system: small changes in concentration would cause large changes in viscosity. In summary: without building in specific interactions neither system usually offers the ideal flow profile. This has led to the use of highly functionalised systems.

#### 11.4.2 Networks and functionalisation

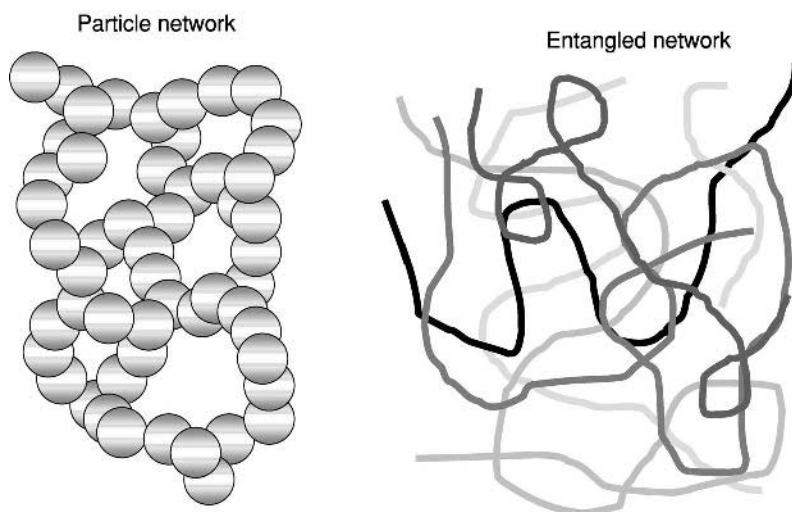
Functionalisation of polymer and particulate systems is very important for achieving good rheology modifiers [3–4]. Manufacturers are constantly striving to develop better additives



**Figure 11.15** The flow curve for a simple entangled polymer as a function of polymer concentration.



**Figure 11.16** The flow curve of a dispersion as a function of volume fraction.



**Figure 11.17** A schematic of the different ways in which connectivity is achieved and networks can be formed.

with lower environmental impact and greater tunability. The basic idea behind the functionalisation of materials is usually twofold:

- to improve the compatibility of the additive to a particular class of materials or commercial sector
- to improve the ‘tunability’ of the rheology through particle–particle or polymer–polymer interactions.

We will find that it is almost universally the case that the rheological properties are determined by controlling the connectivity between the polymers or particles or whatever species we might be using as illustrated in Figure 11.17. In order to get a more viscous or elastic structure we need to have a stronger network. The functionalisation controls how this network is assembled, its strength and what conditions will cause its collapse.

The range of this functionality is vast and it is rare that one additive will uniquely solve the questions we ask of our formulation. The most typical form of functionalisation is the introduction of charge interactions between species. This is most commonly, although not exclusively, used in aqueous systems.

### 11.4.3 Polymeric additives

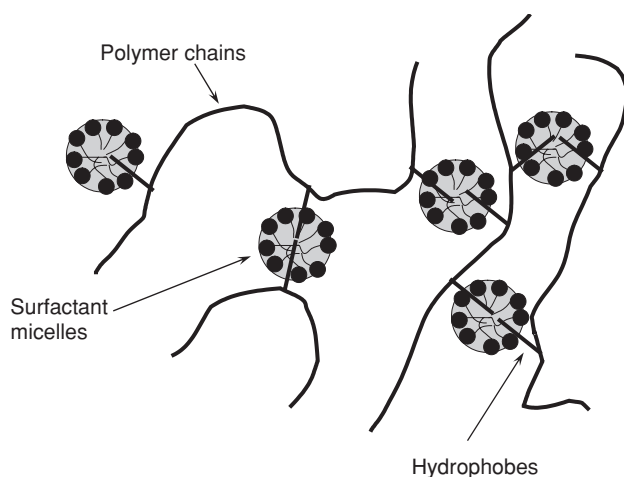
Copolymerisation of polymers can produce block copolymers with charged groups down their backbone. The introduction of acrylic acid for example produces carboxylic acid groups down the backbone of the polymer. This causes elements of the chain to repel one another and the chain to expand. This polymer achieves a higher viscosity at lower concentrations than its non-functionalised counterpart. Of course it increases its sensitivity to both pH and electrolyte. This can be both a problem and an opportunity. It allows the chain expansion and entanglements to be tuned to an extent. At low pH the carboxylic acid groups are relatively

poorly dissociated, leading to lower viscosity systems. One of the difficulties of using such materials is the care with which they must be introduced to our formulation. If exposed to the wrong environment during processing they can precipitate and lose much of their functionality. This can be a 'one way trip' and despite adjustment to the optimum conditions later in the process path, the precipitated material may prove to remain stubbornly in the precipitated state.

It is possible to build network structures by using multi-valent cations. For example the introduction of calcium cations to a polymer rich in carboxylic acid groups can result in ion bridging. The divalent ion can link together charged groups on the same chain or on neighbouring chains. This can even lead to precipitation of the polymer or network formation depending on the relative concentrations and the architecture of the chain. Blends of natural polymers with opposing charges such as acacia and gelatine are a good example where networks can form and modify the rheology but in the correct ratio and at the correct pH they will precipitate to form a coacervate: a solid precipitate. It is possible to produce ionomers where a largely non-aqueous backbone has grafted charge groups. These sites can be used to bridge chains together in low dielectric media since they prefer to associate with each other than expose their charges to media where their opportunity for dissociation is highly limited.

One of the characteristics of particles that is hard to reproduce in polymers is their rapid shear thinning properties. However these characteristics can be achieved using hydrophobic interactions. We can modify an aqueous polymer with hydrophobic groups. For example we can place small chain length alkyl groups as terminating species at either end of a polyethylene oxide chain or as side groups on a hydroxycellulose backbone. These groups prefer to associate with each other rather than the medium or the chain. A schematic of this type of interaction is shown in Figure 11.18, in this case with micelles associating with network sites and the hydrophobic branches of the chain.

The forces holding these sites together are relatively weak. The alkyl chains are attracted by weak van der Waals interactions and the water molecules form a hydrogen bonded cage



**Figure 11.18** A schematic of the role played by hydrophobic modification in forming polymer networks. In this case surfactant is also shown as present in the system.

around them. When we place such a networking point under shear the structure is relatively easy to disrupt to give ‘particle-like’ shear thinning. Hydrophobic modification has been achieved in quite complex molecules such as branched polyacrylamides to give tunable flow behaviour. The alkyl tails are sensitive to surfactants and alcohols which can be used to tune the strength of associations between groups. They provide an interesting range of possibilities.

#### 11.4.4 Particle additives

Particles almost cannot help being functionalised. In general a great deal of care must be taken to produce a dispersion of particles which do not either mutually attract or repel. In polymer latex systems, fragments of the initiator can result in particle surfaces possessing charged groups. This enhances the interactions between the particles producing more viscous systems at lower concentrations than systems interacting through hydrodynamic forces alone.

Colloidal particles are in constant Brownian motion in quiescent conditions. The onset of shear thinning in pseudoplastic dispersions is caused by the shear field that is applied, overcoming the Brownian motion of the particles. The particles are forced to follow the flow lines induced by the movement of the geometry. If we could predict where this occurred we could design materials to flow at a rate appropriate to our needs. In Section 11.2.4 we defined a relaxation time characteristic for the structure. Suppose we consider this time relative to our applied shear rate. The dimension of shear rate is  $s^{-1}$ . So the reciprocal of the rate can be thought of as a characteristic time for the shearing process. If the product of the shear rate and the relaxation time is greater than 1, the convective forces will be stronger than the Brownian forces. If it is less than 1 Brownian forces win out. This ratio is termed the Peclet number,  $Pe$ , and it is given by [3]

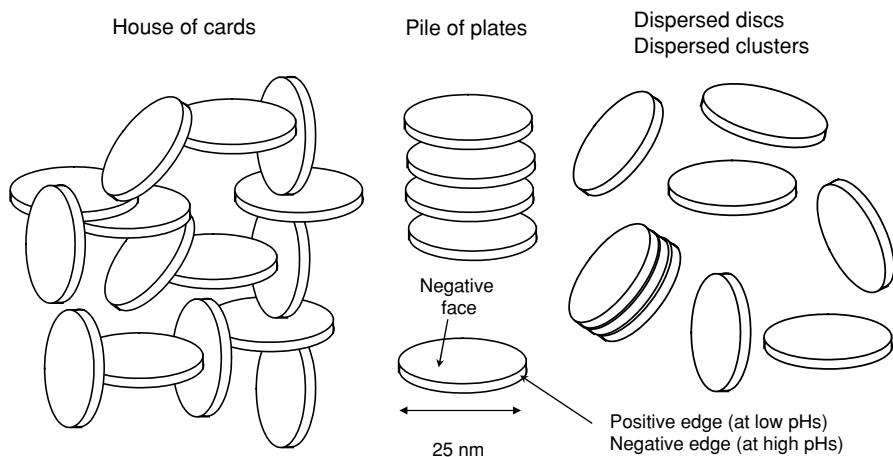
$$Pe = \dot{\gamma}\tau = \frac{6\pi\eta(0)\dot{\gamma}a^3}{kT} = \frac{6\pi\sigma a^3}{kT} \quad (11.20)$$

The Peclet number defines the stress where we would start to see shear thinning. So if we set the Peclet number equal to 1 we can define a relationship between the critical stress for shear thinning,  $\sigma_c$ , the particle radius and temperature:

$$\sigma_c = \frac{kT}{6\pi a^3} \quad (11.21)$$

From this we can see that the smaller the particle the higher the stress we require to cause it to undergo shear thinning. In order to disrupt this relationship we need to impart additional forces to the particles.

Attractive forces between particles can be controlled to provide weak interactions and the tendency to weakly aggregate. This can aid storage stability. At first it seems counterintuitive to use aggregation to control this type of stability. We might suppose as particles aggregate they become larger and thus more prone to separate. However if the concentration is high enough it can lead to fairly stable networks. These become space filling at much lower concentrations than 60%. They can develop self-supporting network structures that resist sedimentation and that are highly shear thinning. As we saw in Section 11.3.5, frequently we only need a weak yield stress to oppose the effects of gravitational forces. This can be



**Figure 11.19** An example of some of the ways in which Laponite discs may associate.

achieved with relatively weak attractions between the particles. The application of even a small shearing stress readily disrupts this network and causes substantial flow, and hence shear thinning to occur. Thus systems which aggregate weakly can be used to build networks stable to sedimentation but fragile at greater shear stresses.

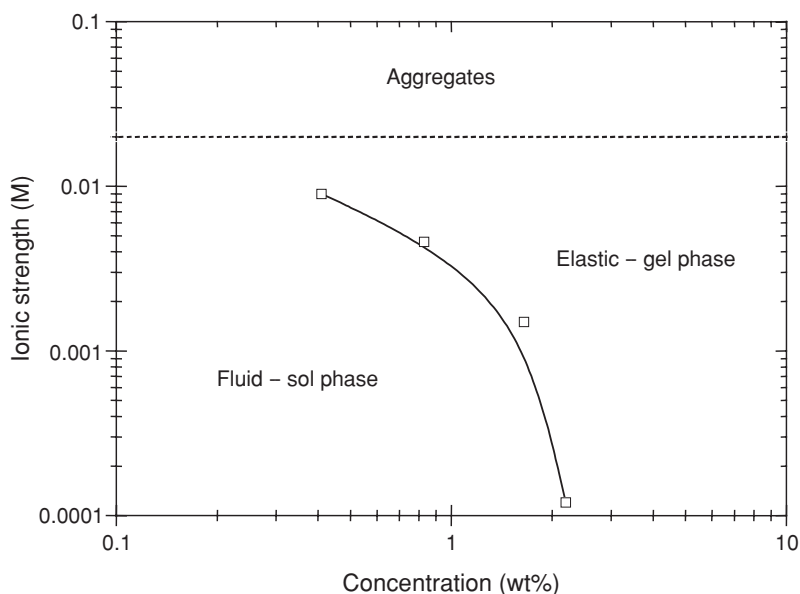
So far we have restricted our thinking to spherical colloidal particles. The most catastrophic shear thinning is achieved with particles with strong shape anisotropy. Rod-like or disc-like particles can show tremendous changes in rheology. For example the synthetic hectorite clay, Laponite, is a small disc-like entity. The edges and the faces can take different charges depending on the pH and ionic strength of the solution. They can form clusters, they can mutually repel or perhaps form edge-face house-of-cards-like structures as shown in Figure 11.19.

At high pH both the faces and edges carry the same sign of electrical charge and mutual repulsion results at low levels of ionic strength. This forms a gel-like material at high particle concentrations and a pseudoplastic fluid at low particle concentrations. As the electrolyte is increased the system begins to flocculate leading to weak yield stresses. This material is highly tunable in the appropriate formulation. The phase diagram for transitions at high pH is outlined in Figure 11.20.

Laponite is just one example of a whole range of clays. It is particularly useful in 'clean' applications such as pharmaceutical and personal care products. Organophilic clays possess some of the flexibility of their aqueous relations although they tend to be less tunable.

A simple trick we can play with particulate systems is to increase their porosity, although the word simple should be viewed advisedly. Porous particles, especially those that are siliceous, tend to form highly space filling structures when dispersed into a formulation. The maximum packing fraction occurs at very low mass concentrations and a small amount of additive can result in large rheological changes. These systems are guilty of having large surface areas. They can both adsorb species and leach ions leading to temporal drifts in properties in the wrong formulation. However there is a wide variety of choices of materials to select from. They can often be tailored to produce the required behaviour.





**Figure 11.20** A phase diagram for a Laponite dispersion at a high pH.

## 11.5 Summary

In this chapter we have examined some of the important features of experimental design, methods and the chemical influences on the rheological properties of colloidal systems. This treatise has indicated problems to be aware of when making measurements in order to avoid artefact laden data. Our choice of experiment should be such that it is 'fit for purpose' and is either related to our process or can be used to give an interpretation of the structural properties of our system. Finally, we have touched upon the behaviour of polymers and particles and how their properties relate to their rheology. It is difficult to do full justice to the field of experimental rheology in this chapter. The reader is urged to examine the referenced texts and be rewarded by greater insights into specific aspects of the subject.

## References

1. Macosko, C.W. (1994) *Principles, Measurements and Applications*. Wiley-VCH, New York.
2. Bird, R.B. Stewart, W.E. and Lightfoot, E.N. (1960) *Transport Phenomena*, Wiley, New York.
3. Goodwin, J.W. and Hughes, R.W. (2000) *Rheology for Chemists, An Introduction*. The Royal Society of Chemistry, Cambridge.
4. Larson, R.G. (1999) *The Structure and Rheology of Complex Fluids*. Oxford University Press, Oxford.
5. Ferry, J.D. (1989) *Viscoelastic Properties of Polymers*. Wiley, New York.
6. Doi, M. and Edwards, S.F. (1986) *The Theory of Polymer Dynamics*. Oxford University Press, Oxford.
7. Gross, B. (1968) *Mathematical Structures of the Theories of Viscoelasticity*. Hermann, Paris.
8. Tschoegl, N.W. (1989) *The Phenomenological Theory of Linear Viscoelastic Behaviour*. Springer-Verlag, Berlin.

## Chapter 12

# Scattering and Reflection Techniques

*Robert Richardson*

### 12.1 Introduction

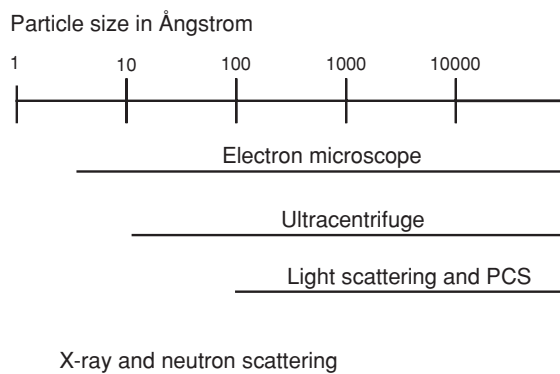
Scattering of radiation is an essential tool for the colloid scientist. This chapter aims to introduce the use of scattering techniques in colloid science. For more detailed information, there are several books that are based on the science that can be done using particular types of radiation [1–3]. There are also several books that describe the structure determination for crystals and glasses where the focus is the arrangement of atoms rather than the larger scale structures of interest to colloid scientists. Nevertheless, these are useful resources for information of relevance to colloidal systems [4]. The most useful books for a colloid scientist seeking to consolidate this introduction are those that describe the scattering methods applied to colloids and polymers [5].

Colloids are generally dispersions of particles with dimensions between 1 nm and 1  $\mu\text{m}$ , i.e. between 10 and 10 000 Å (since 10 Å = 1 nm). There are many experimental techniques for characterising particles in this size range and their approximate ranges of sensitivity are shown in Figure 12.1 and the strengths and weaknesses of the different techniques are discussed below.

Electron microscopy (discussed in Chapter 13) can cover the whole range of sizes and gives very detailed results. The only limitation is that in general the sample is not in its natural equilibrium state. A vacuum is generally necessary so a colloidal dispersion would be ‘frozen’ or its particles extracted and dried for ‘*ex situ*’ electron microscopy. To overcome these limitations, there are continuing developments in this field and environmental scanning electron microscopy is capable of operating under a liquid vapour while cryo-transmission electron microscopy seeks to freeze in an equilibrium structure by very rapid cooling.

There are several techniques for particle size determination based on sedimentation. These give useful but rather limited information on particles in an equilibrium state. An example is the ultracentrifuge that can cover a wide range of particle size but experiments are quite demanding. Simple measurements of solution viscosity can also be used to estimate particle size.

Light scattering and small angle X-ray and neutron scattering by the particles in a dispersion are the main topic of this chapter. These scattering techniques cover the entire size range of interest and are capable of measuring the dimensions of particles ‘*in situ*’. These methods may also give more detailed information on the internal structure of particles and their interactions in a dispersion. It is this ability to determine the equilibrium structure



N.B.  $1 \text{ Å} = 0.1 \text{ nm}$

**Figure 12.1** Size ranges covered by different sizing techniques.

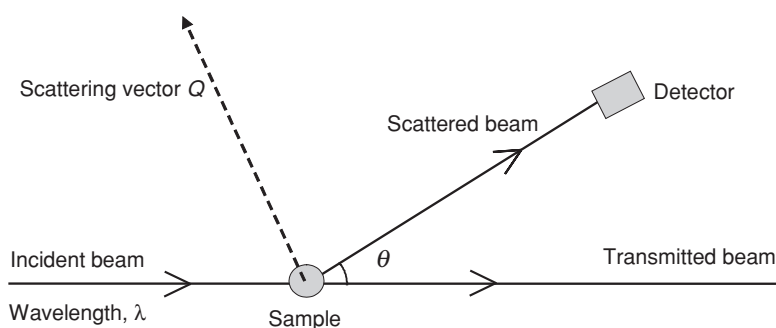
of colloids and surfaces in detail using a non-invasive probe that gives scattering methods importance in colloid and surface science.

## 12.2 The Principle of a Scattering Experiment

The basic scattering experiment is very simple. As shown in Figure 12.2, a monochromatic (i.e. single wavelength) beam is brought onto a sample.

The intensity of the scattered radiation is measured as a function of the scattering angle which we will label as  $\theta$  but note there are other conventions. The important variable, however, is the scattering vector,  $Q$ , whose magnitude is related to the scattering angle and wavelength:

$$Q = \frac{4\pi \sin \theta/2}{\lambda} \quad (12.1)$$



**Figure 12.2** Schematic diagram of a scattering experiment.

In principle, one could measure some scattering with two different wavelengths and plot the intensity versus  $Q$  and get the same features of the curves at the same  $Q$  values. The distances probed in an experiment are inversely proportional to  $Q$  (i.e. distance  $\sim 2\pi/Q$ ). That means for large scale structures (e.g. 100 Å or 10 nm) we need a small  $Q$  (i.e.  $Q \sim 0.06 \text{ Å}^{-1}$ ). To achieve small  $Q$  in a scattering experiment we need a suitable combination of large wavelength and small scattering angle. For light scattering, a wavelength that is comparable to or larger than the size of the scattering particles is generally used. For X-ray and neutron scattering a low scattering angle is generally used.

## 12.3 Radiation for Scattering Experiments

Table 12.1 summarises the properties of visible light, X-rays and neutrons for scattering experiments from colloidal dispersions.

Visible light has a wavelength of 400–600 nm and is suited to particle sizes above  $0.01 \mu\text{m}$  although smaller particles can be detected. It is scattered by particles that have a different refractive index from that of the surrounding solvent.

X-rays (like visible light) are electromagnetic radiation. The useful wavelength range is shorter than about 0.2 nm because longer wavelengths tend to be adsorbed strongly. Small angle X-ray scattering is suited to probing distances in the 1 nm– $1 \mu\text{m}$  range. For X-rays, it is the difference in mean electron density between a particle and the solvent that scatters the radiation. They are therefore excellent for dispersions of high atomic number materials in low atomic number solvents (e.g. metals or oxides in water). They are less good for dispersions of organic materials in aqueous solvents because the electron densities of the two materials are similar.

Neutrons are particles but have an associated wavelength. The useful wavelength range is about 0.1–2 nm. Longer wavelength neutrons are difficult to produce with sufficient intensity. Small angle neutron scattering is suited to probing distances in the 1 nm– $1 \mu\text{m}$  range. Neutrons do not generally interact with the electrons in atoms but they are scattered by an interaction with the nuclei. The scattering is not dependent on the atomic number, so they often offer a better option than X-rays for scattering from a dispersion of one material in another of similar atomic number. It is also possible to get different scattering from the same element by using isotopes. Hydrogen/deuterium substitution is particularly useful in colloid science. For instance, organics dispersed in heavy water ( $\text{D}_2\text{O}$ ) scatter neutrons strongly. First we will look at the factors that determine the intensity in a classic light scattering experiment.

**Table 12.1** Properties of radiation for scattering experiments

Radiation	Visible light	X-rays	Neutrons
Type	Electromagnetic wave	Electromagnetic wave	Particle/wave
Wavelength, $\lambda$	400–600 nm	0.01–0.2 nm	0.01–2.0 nm
Distances probed	$>0.01 \mu\text{m}$	nm to $\mu\text{m}$	nm to $\mu\text{m}$
Scattered by	Refractive index	Electron density	Nuclear scattering
variations of			properties

## 12.4 Light Scattering

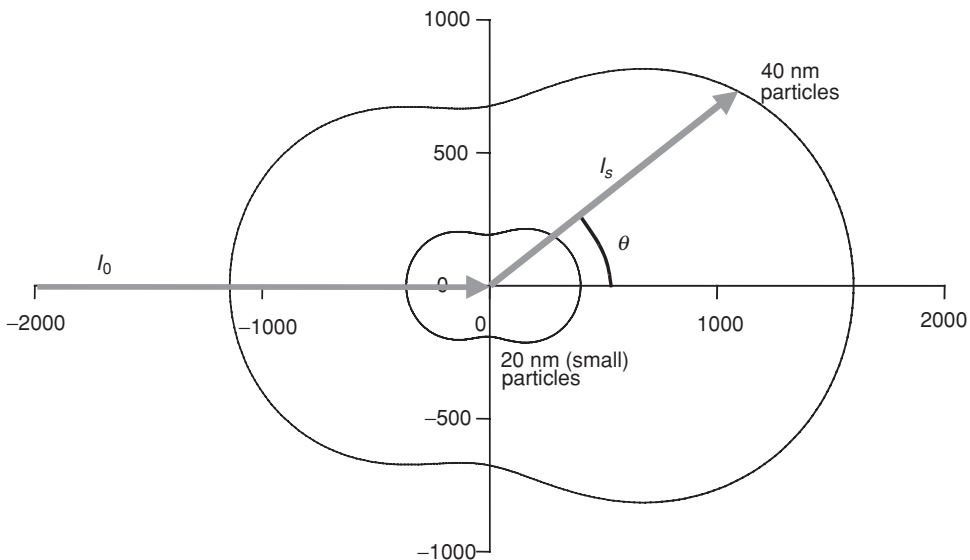
Light scattering has been used for decades to determine particle sizes. The intensity of light scattered by a suspension of 'small' particles (i.e. particle diameter  $\ll$  wavelength) is determined by

$$I(Q) = kcM(1 + \cos^2 \theta) \quad (12.2)$$

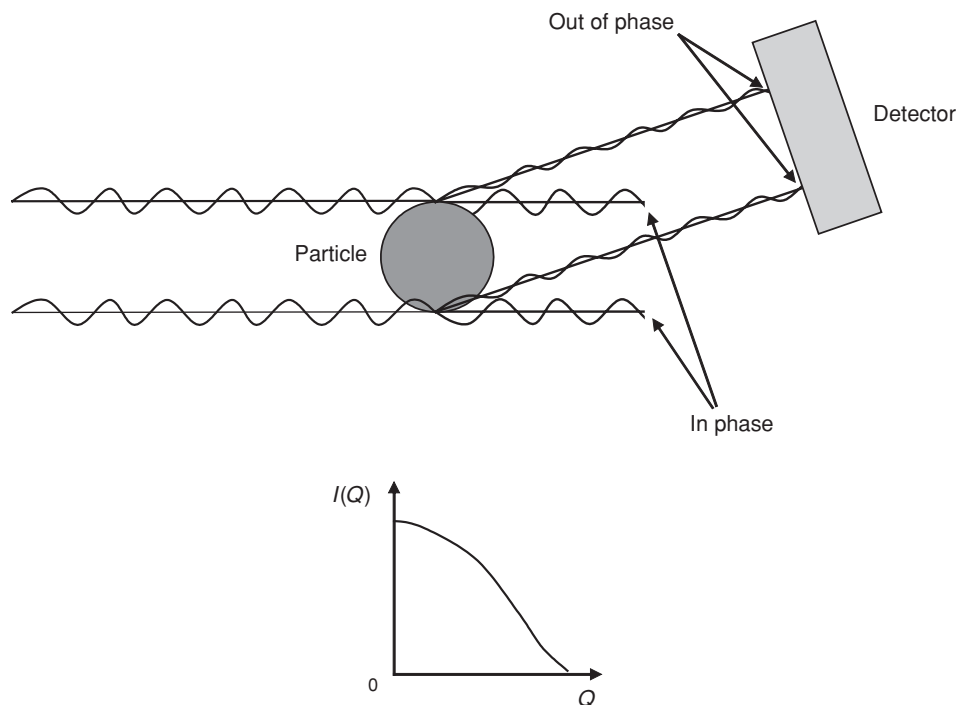
where the four factors are the concentration,  $c$ , the molar mass of the particles,  $M$ , a collection of constants,  $k$ , and the polarisation factor  $(1 + \cos^2 \theta)$ . The polarization factor results from the physics of the scattering process and is usually removed by an experimental correction. It contains very little information about the sample. This equation was derived by Lord Rayleigh in 1871. An interesting sideline is that the factor  $k$  contains the inverse fourth power of the light wavelength which means that short wavelengths are scattered much more strongly than long ones. Hence the sky is blue and the sun appears yellow or red.

When light of wavelength 500 nm is incident on small particles (20 nm radius) the anisotropy of the scattering only results from the polarisation factor. Forward and backscattering intensities are therefore equal as shown in the polar plot of scattered intensity in Figure 12.3. If the constant,  $k$ , and concentration,  $c$ , are known the particle mass may be determined.

However, for larger particles (40 nm) the scattered intensity develops an asymmetry between forward and backward scattering. This is a particle size effect and is the basis for the determination of particle size from scattering. The origin of this effect is shown in Figure 12.4. If a particle radius is comparable with the wavelength of the radiation, there will be a different path length from source to detector for rays scattered by different parts



**Figure 12.3** Polar plot of forward and backscattering intensities.



**Figure 12.4** Scattering from a 40 nm particle.

of the particle. Consider two rays scattered from opposite sides of the particle. They have different source to detector distances. This path difference means that the two rays arrive at the detector somewhat out of phase. They interfere destructively so there is less intensity detected. In fact the degree of destructive interference tends to increase as the scattering angle increases so the intensity tends to decrease with increasing angle or  $Q$ .

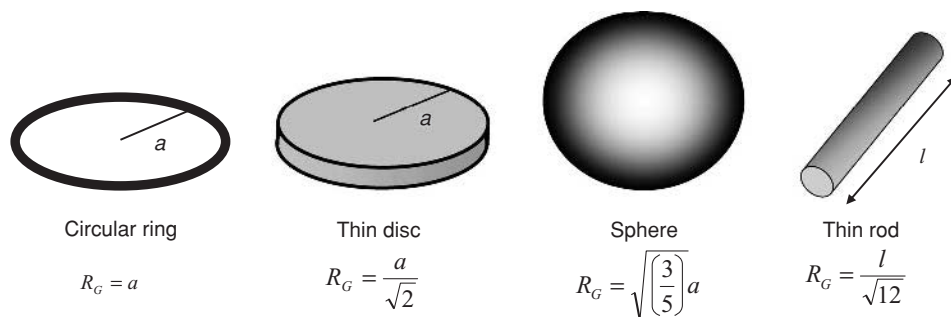
The rate at which the intensity changes with angle (or  $Q$ ) depends on the size of the particle. However, 'the size of a particle' is a vague concept. More precisely, the intensity depends on the radius of gyration of the particle. The radius of gyration,  $R_G$ , depends on the dimensions and shape of a particle. In general it may be calculated for any shape by an integral over the volume of a single particle:

$$R_G = \sqrt{\frac{1}{V} \int_V r^2 dV} \quad (12.3)$$

where  $r$  is the distance from the centre of gravity and  $V$  is the volume of the particle. The relationship of  $R_G$  to the dimensions of some simple shapes is given in Figure 12.5.

For large particles, there is an additional factor in the equation that governs the scattered intensity:

$$I(Q) = k c M (1 + \cos^2 \theta) (1 - (Q R_G)^2/3 + \dots) \quad (12.4)$$



**Figure 12.5** Some simple particle shapes and their radii of gyration.

This new factor depends on the radius of gyration and is approximated by  $(1 - (QR_G)^2/3)$ . Hence the dependence of the scattered intensity on scattering angle or  $Q$  may be used to determine the radius of gyration of particles in a dilute dispersion. This theory (known as the Rayleigh–Gans–Debye theory) is applicable where particle diameter remains less than the wavelength. For larger particles the scattering pattern becomes extremely complex and Mie theory applies. This is covered in more advanced texts such as [3].

## 12.5 Dynamic Light Scattering

There is another light scattering technique that operates in a different way. It relies on the fact that particles in a dispersion are moving by diffusion. When a particle scatters a photon of light, there is a small exchange of energy between the photon and the particle. The particle may gain energy from or lose energy to the photon and the photon's energy shifts accordingly. This is the same process (Doppler shifting of frequency) that is used in radar speed traps. The frequency of the radar is changed by reflection from a moving vehicle.

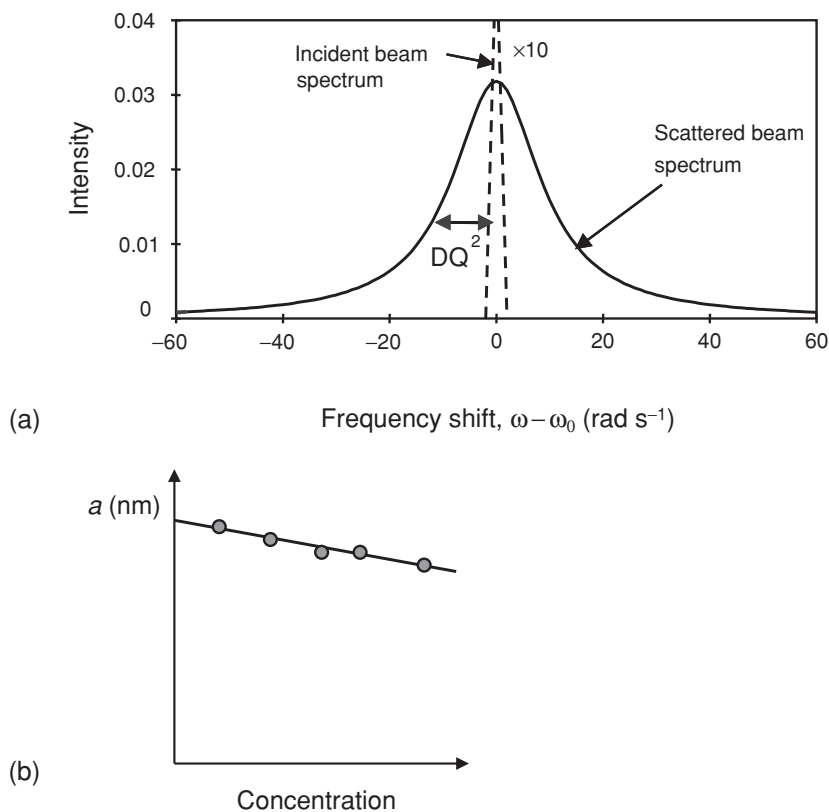
The spectrum of light scattered is measured using the elegant technique of photon correlation spectroscopy (PCS). If the incident spectrum is monochromatic with frequency,  $\omega_0$ , then the spectrum from a colloid generally has a Lorentzian shape and the width of the peak is determined by the diffusion coefficient of the particles,  $D$ , multiplied by  $Q^2$  as shown in Figure 12.6:

$$I(\omega) \propto \frac{DQ^2}{(\omega - \omega_0)^2 + DQ^2} \quad (12.5)$$

The hydrodynamic radius,  $a$ , may then be calculated from the diffusion coefficient using the Stokes–Einstein equation, provided the solvent viscosity,  $\eta$ , is known (more details may be found in [6]):

$$a = \frac{kT}{6\pi\eta D} \quad (12.6)$$

Since the diffusion coefficient is influenced by particle–particle interactions as well as viscous drag, it is usually necessary to extrapolate the hydrodynamic radius to zero concentration.



**Figure 12.6** (a) Incident and scattered spectra for dynamic light scattering and (b) extrapolating the hydrodynamic radius to zero concentration.

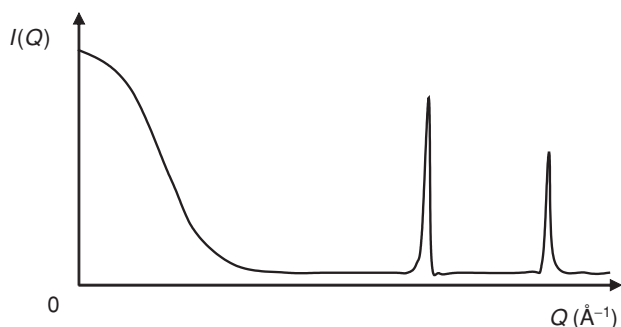
It should be noted that the hydrodynamic radius is often greater than the radius determined by 'static' light scattering because it may include layers of bound solvent.

## 12.6 Small Angle Scattering

We now consider the techniques of small angle scattering of X-rays and neutrons (SAXS and SANS). A better term would be small  $Q$  scattering because the exact combination of scattering angle and wavelength is not important.

Take a gold colloid for example. The scattering from such a sample is shown schematically in Figure 12.7. At small  $Q$ , the scattering is sensitive to the size and shape of the particles (as we saw for light scattering) while at large  $Q$  the scattering reflects the internal structure of the particles. For particles of a crystalline material the internal structure would give Bragg peaks. The Bragg peaks from colloidal particles are often broader than those from the same material in bulk form because of internal disorder and finite size effects on the diffraction. However it is the small  $Q$  scattering that is generally measured and interpreted in colloid science.





**Figure 12.7** Schematic of scattering from colloid of crystalline particles.

## 12.7 Sources of Radiation

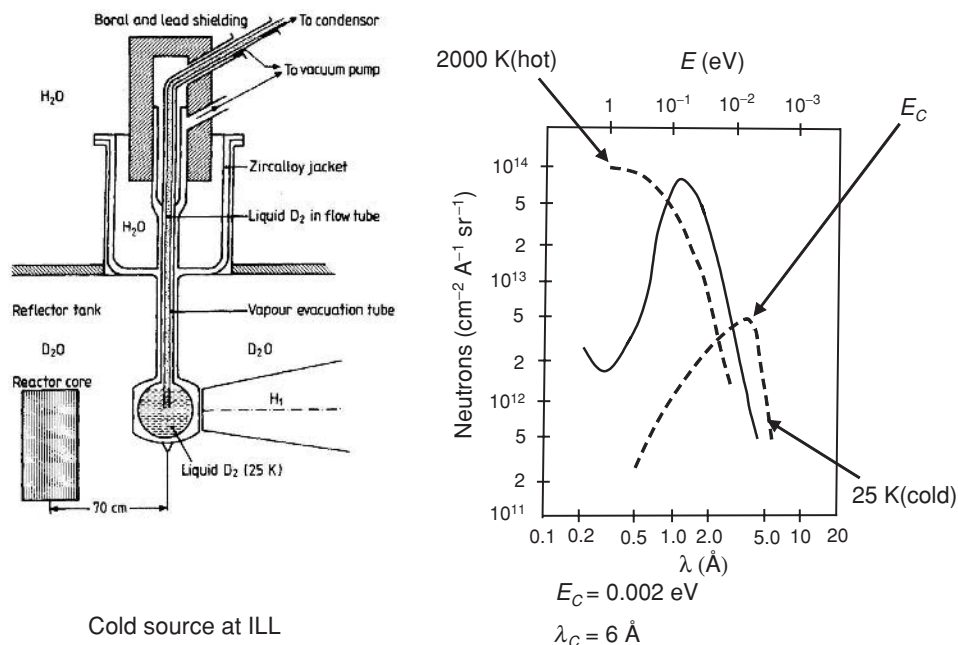
X-rays may be generated in the laboratory using a sealed tube generator. A complete SAXS kit can be bought from several manufacturers. The intensity from a synchrotron source is many orders of magnitude greater. Synchrotron radiation is produced tangentially when a high energy electron beam is deflected by a magnetic field. Facilities exist at the Daresbury SRS, the ESRF at Grenoble and Diamond is under construction at the Rutherford Appleton Laboratory, Oxfordshire. More details may be found at the facility Web sites [7–9]. Neutrons in adequate quantities are only available from central facilities. In the United Kingdom we have convenient access to the reactor at ILL, Grenoble [10] and the pulsed source at ISIS at the Rutherford Appleton Laboratory [11]. A second source (Target Station 2) is under construction at ISIS. Further details and links to neutron sources worldwide may be found at the facility Web sites. The European Neutron Scattering Association gives information for accessing European Facilities and is actively involved in planning future sources [12]. Neutrons produced by the fission process in a reactor have a high energy ( $E \sim 1$  MeV) and by consequence of the de Broglie relationship,

$$\lambda(\text{\AA}) = 9.04/\sqrt{E(\text{meV})} \quad (12.7)$$

a very short wavelength ( $\lambda \sim 0.0003$  Å). This is useless for large scale structures. Fortunately the energy of neutrons may be moderated by passing them through a material to thermalise. The neutrons adopt the thermal energy of the moderator material. If its temperature is low, the energy of the neutrons becomes low and so their wavelength becomes large. At the ILL, the cold moderator is liquid deuterium at 25 K and it turns out that a 25 K liquid deuterium moderator gives a ‘Maxwellian’ distribution of wavelengths peaked around 6 Å. Such a cold source is ideal for SANS. Figure 12.8 shows the moderator schematically and its wavelength distribution along with those from ambient and hot moderators. Pulsed neutron sources also use moderators to generate neutrons in the useful wavelength range.

## 12.8 Small Angle Scattering Apparatus

The basic components of a small angle neutron scattering apparatus (SAXS is similar after the sample) are shown in Figure 12.9. The reactor core is surrounded by D<sub>2</sub>O to reflect

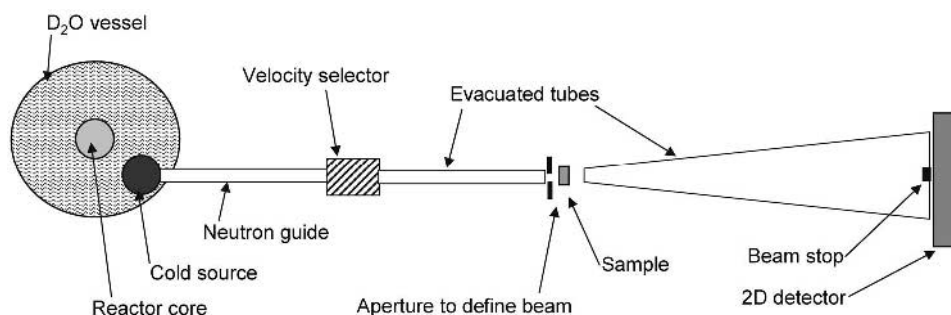


**Figure 12.8** Schematic of a neutron moderator and the resultant wavelength distribution. Reprinted from Bee, M. (1988) *Quasielastic Neutron Scattering*. Institute of Physics Publishing Ltd, Bristol, with permission.

neutrons back and maximise flux. The cold source is used to maximise the useful flux at  $\sim 6 \text{ \AA}$ . The velocity selector only passes a narrow(ish) band of neutron velocities. Since velocity,  $v$ , and wavelength,  $\lambda$ , are inversely proportional,

$$\lambda(\text{\AA}) = 3096/(v(\text{ms}^{-1})) \quad (12.8)$$

a narrow band of wavelengths is passed and so the radiation incident on the sample is nearly monochromatic. The sample is typically  $1 \text{ cm}^2$  and  $1 \text{ mm}$  thick. A two dimensional position sensitive detector collects scattered intensity. Software is often used to regroup the data as intensity versus  $Q$ . Evacuated tubes reduce background from air scattering.



**Figure 12.9** Schematic diagram of a small angle neutron scattering apparatus.

On a pulsed neutron source, such as ISIS, the velocity selector is unnecessary since the time of flight or velocity of the neutrons can be measured. From this the speed and hence wavelength can be calculated. Since a ‘white beam’ is used the large intensity losses associated with monochromatisation are avoided. Although intrinsically weaker, pulsed sources tend to make efficient use of the neutrons.

Figure 12.10 shows the NG3 SANS apparatus at NIST. D11 and D22, the SANS instruments at ILL, have similar layout. Other reactor based facilities also have similar instrumentation. The SANS apparatus at ISIS is LOQ and more instruments are to be built on the second target station.

## 12.9 Scattering and Absorption by Atoms

The amplitude of the scattering by an atom is characterised by its scattering length,  $b$ . For X-ray scattering,  $b$  is proportional to the atomic number,  $z$ . (Actually  $b = z \times a_e$  where  $a_e = 2.85 \times 10^{-15}$  m, which is the scattering length for one electron.) For neutrons, the scattering length is a nuclear property and it varies irregularly with atomic number and also depends on isotope. Table 12.2 shows the scattering lengths and absorption cross sections for some atoms. We see that hydrogen and deuterium have very different scattering lengths. Physically, a positive or negative scattering length is related to the phase shift of the wave during scattering but we do not need to be concerned over the origin of the sign. We just use it.

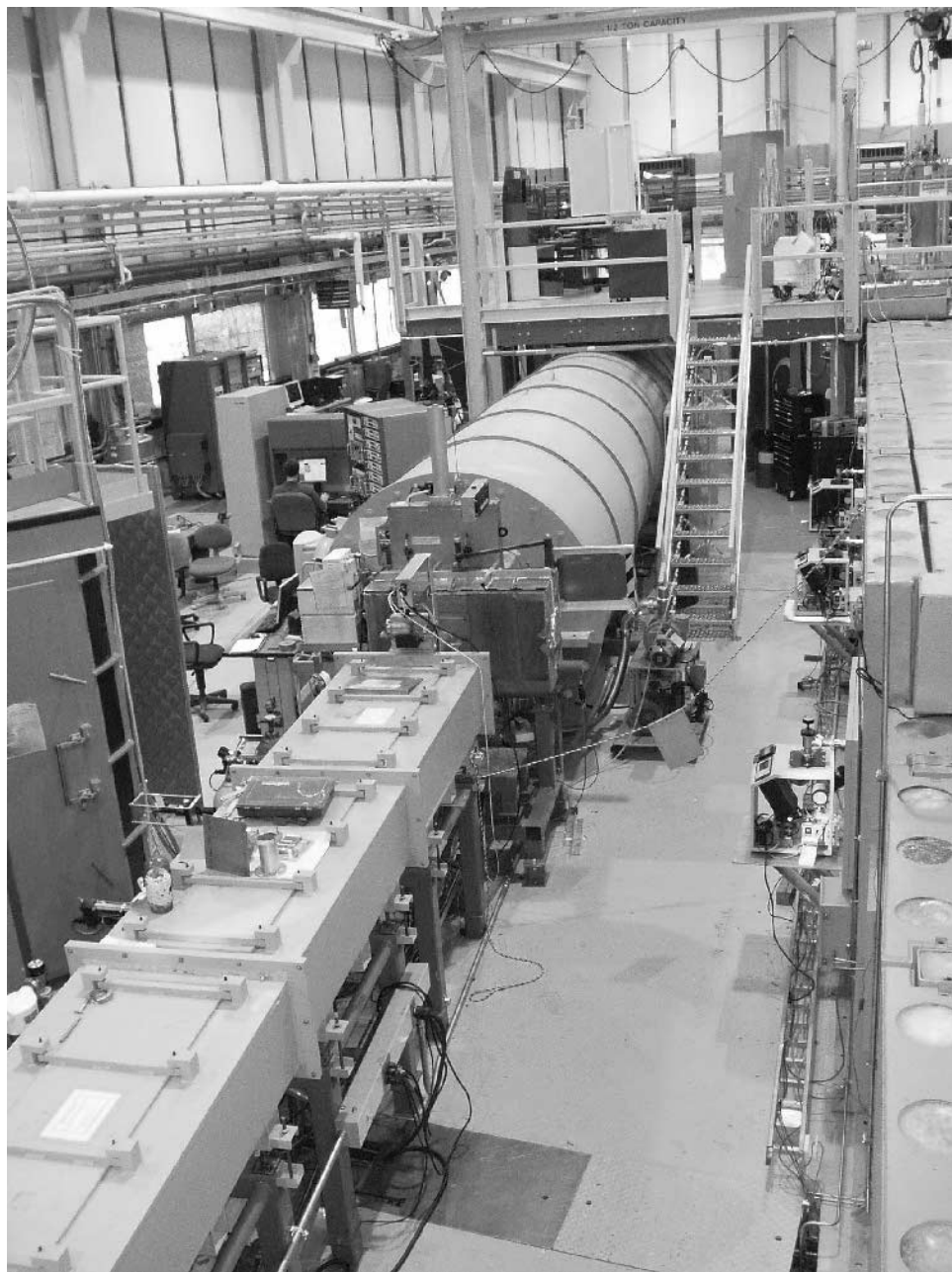
The absorption cross sections indicate how strongly an element absorbs the radiation. Absorption of X-rays increases very strongly with atomic number, so cells for X-ray experiments are made of low- $z$  materials. Neutrons tend to be absorbed less, so sample containment is not a problem. There are useful exceptions such as cadmium which can be used for shielding and beam definition apertures.

### 12.10 Scattering Length Density

In ‘small angle’ experiments (i.e. low  $Q$ ) the distances probed are generally much greater than inter-atomic spacings, so the technique is sensitive to changes in ‘scattering length density’ over distances of up to  $\sim 1000$  Å rather than the scattering by individual atoms. Scattering length density,  $\rho$ , is therefore a very useful concept because it can be used to

**Table 12.2** Some scattering lengths and absorption cross sections [13]

Species	$b_N 10^{-15}$ m	$b_X(\text{at } Q = 0) 10^{-15}$ m	$\sigma_N(\text{abs}) 10^{-28} \text{ m}^2$	$\sigma_X(\text{abs}) 10^{-28} \text{ m}^2$
H	−3.74	2.85	0.28	0.73
D	6.67	2.85	0.0	0.73
C	6.65	17.1	0.003	92
O	5.83	22.8	0.0	306
Cd <sup>2+</sup>	3.7	131.1	>103	9400



**Figure 12.10** The NG3 small angle neutron scattering apparatus at the NIST Centre for Neutron Research, Gaithersburg, MD, USA. It shows the shielding around the incident beam in the foreground, the sample position and the large vacuum tank containing the detector in the background.

**Table 12.3** Scattering length densities of some materials

Material	$\rho_N$ ( $10^{-5} \text{ \AA}^{-2}$ )	$\rho_X$ ( $10^{-5} \text{ \AA}^{-2}$ )
H <sub>2</sub> O	-0.05	0.94
D <sub>2</sub> O	0.64	0.94
-(CH <sub>2</sub> ) <sub>n</sub> -	-0.06	0.65
-(CD <sub>2</sub> ) <sub>n</sub> -	0.61	0.65

describe the scattering from a large volume (such as a particle) without having to specify the position of every atom. The scattering length density of a material,  $\rho$ , is calculated from the product of the number density of each atom type,  $N_j$ , and its scattering length,  $b_j$ . The products for different types of atom are then summed.

For neutrons:

$$\rho_N = \sum N_j b_j \quad (12.9)$$

where  $b_j$  is the scattering length of an atom for neutrons.

For X-rays:

$$\rho_X = a_e \sum N_j z_j \quad (12.10)$$

where  $a_e$  is the scattering length of an electron for X-rays ( $a_e = 2.85 \times 10^{-5} \text{ \AA}$ ) and  $z_j$  is the atomic number.

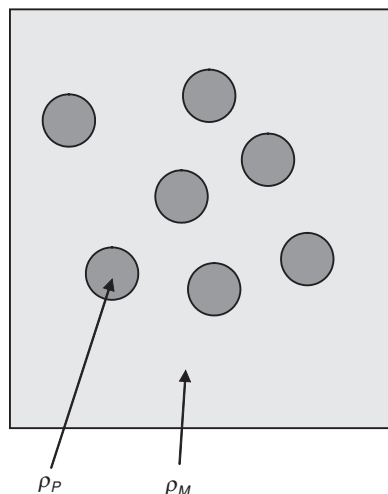
The scattering length densities for some materials are shown in Table 12.3. It is worth noting that for neutrons the scattering length density depends on the isotope. This gives the experimentalist an important tool because it is possible to vary the isotopic content of a sample, for instance by switching from H<sub>2</sub>O to D<sub>2</sub>O as solvent, in order to emphasise some aspect of the scattering without changing the chemistry of the sample appreciably.

## 12.11 Small Angle Scattering from a Dispersion

A simple picture of a dispersion is a number of identical particles suspended in a matrix (the solvent) as shown in Figure 12.11. For a dilute dispersion the inter-particle distance will be able to take almost any value and so inter-particle interference effects are eliminated. The observed scattering intensity,  $I(Q)$ , then depends only on the four factors in the following equation:

$$I(Q) = (\rho_P - \rho_M)^2 N_P V_P^2 P(Q) \quad (12.11)$$

where  $(\rho_P - \rho_M)$  is the *contrast* in scattering length density between a particle and the matrix,  $N_P$  is the number of particles in the sample,  $V_P$  is the volume of a particle and  $P(Q)$  is the particle form factor which is defined by the size and shape of the particle.



**Figure 12.11** Simple picture of a dispersion of homogeneous particles in a matrix.

## 12.12 Form Factor for Spherical Particle

The form factor may be calculated by integration over the volume of a particle of any shape. For a spherical particle, the formula given below is obtained [14]:

$$P(Q) = \left\{ \frac{3(\sin QR_S - QR_S \cos QR_S)}{(QR_S)^3} \right\}^2 \quad (12.12)$$

where  $R_S$  is the radius of the sphere.

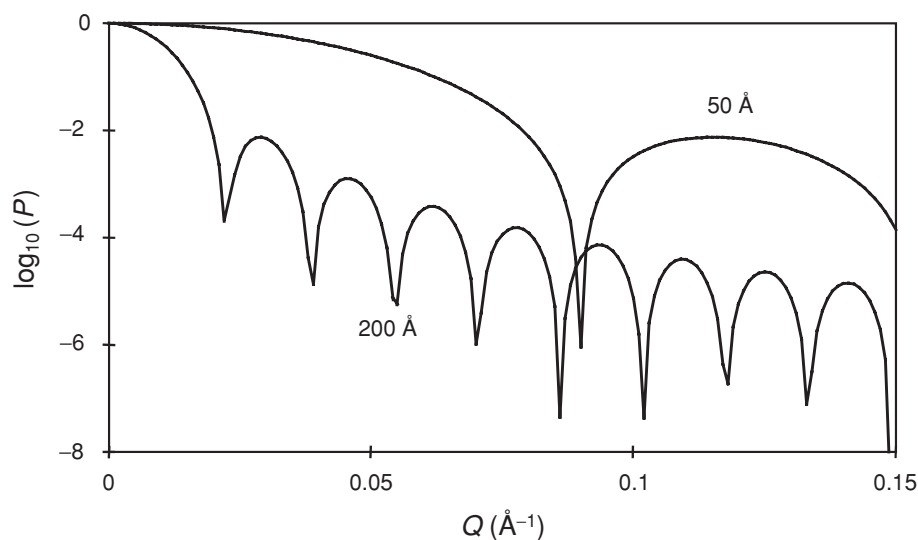
This formula is plotted in Figure 12.12 for two values of the sphere radius. It demonstrates several of the important general features of form factors (NB the log scale)

- At  $Q = 0$  it has a value of 1.
- Initially it decreases with increasing  $Q$ .
- For smaller particles the function is more stretched out in  $Q$  and vice versa. In fact it is a function of  $QR_S$ .
- Maxima and minima appear at higher  $Q$ .

## 12.13 Determining Particle Size from SANS and SAXS

There are two complementary approaches to determining the particle dimensions in a dilute dispersion.

- By calculating the scattering from an assumed particle shape (e.g. a sphere) and varying the parameters (e.g. radius, number of particles) until a good agreement is found between measured data and model. If no agreement is found, then we assume another shape and



**Figure 12.12** Form factor of spherical particle.

repeat the fitting process. This is usually done with a least squares fitting program. It is a useful method but the ‘black-box’ approach may lead to incorrect conclusions regarding the shape.

- The Guinier law relates the low  $Q$  slope of the scattering to the radius of gyration of the particle and makes no assumptions regarding the particle shape.

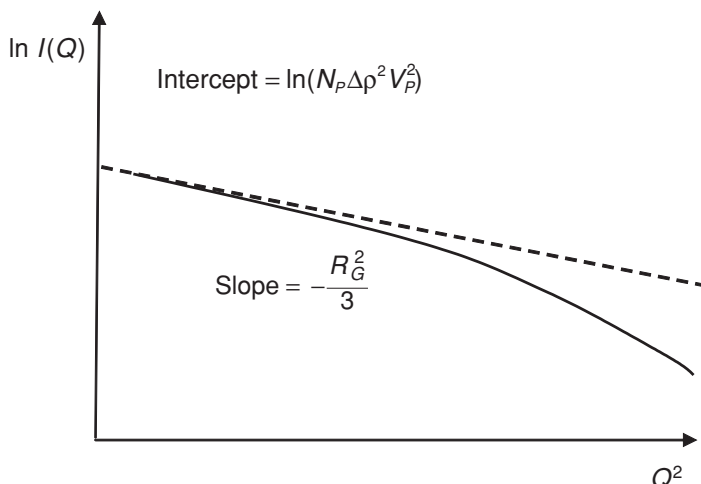
## 12.14 Guinier Plots to Determine Radius of Gyration

It turns out that at low  $Q$  ( $Q < 1/R_G$ ) the scattering from a dilute dispersion is insensitive to the shape of the particles. The intensity,  $I(Q)$ , only depends on contrast, number of particles, particle volume and the radius of gyration as shown in this approximate equation, known as Guinier’s law [15]:

$$I(Q) = \Delta\rho^2 N_P V_P^2 \exp(-Q^2 R_G^2/3) \quad (12.13)$$

The radius of gyration was introduced in Section 12.4 for light scattering and is a very convenient quantity for characterising the size of a particle. Figure 12.13 shows a Guinier plot of natural log of intensity versus  $Q^2$ . It has a slope of  $-R_G^2/3$ . So it is therefore possible to determine  $R_G$  without assuming a shape. Note that caution is required because the approximation is only valid for  $Q < 1/R_G$ .

In the next few sections we look at variations and extensions of this basic type of measurement. Much more detail is available in specialised texts [16, 17].



**Figure 12.13** Showing the straight line behaviour on a Guinier plot at low  $Q$ .

## 12.15 Determination of Particle Shape

At  $Q > 1/R_G$  the shape of the particle does have a major influence on the particle form factor and hence the shape of the scattering from a dilute suspension. This can be seen most clearly on a log-log plot of the particle form factor for a sphere, a thin (i.e. 5 Å thick) disc and a thin (i.e. 5 Å radius) rod as shown in Figure 12.14. This shows a characteristic region with a slope of  $-1$  for the rod and  $-2$  for the discs. At  $Q \gg 1/(\text{the dimension of the particle})$ , Porod's law (discussed later) applies. The particle sizes in this example have been chosen to have the same radius of gyration (100 Å), so the form factor is the same for all three in the region below  $Q \sim 1/R_G$  where the scattering obeys the Guinier law.

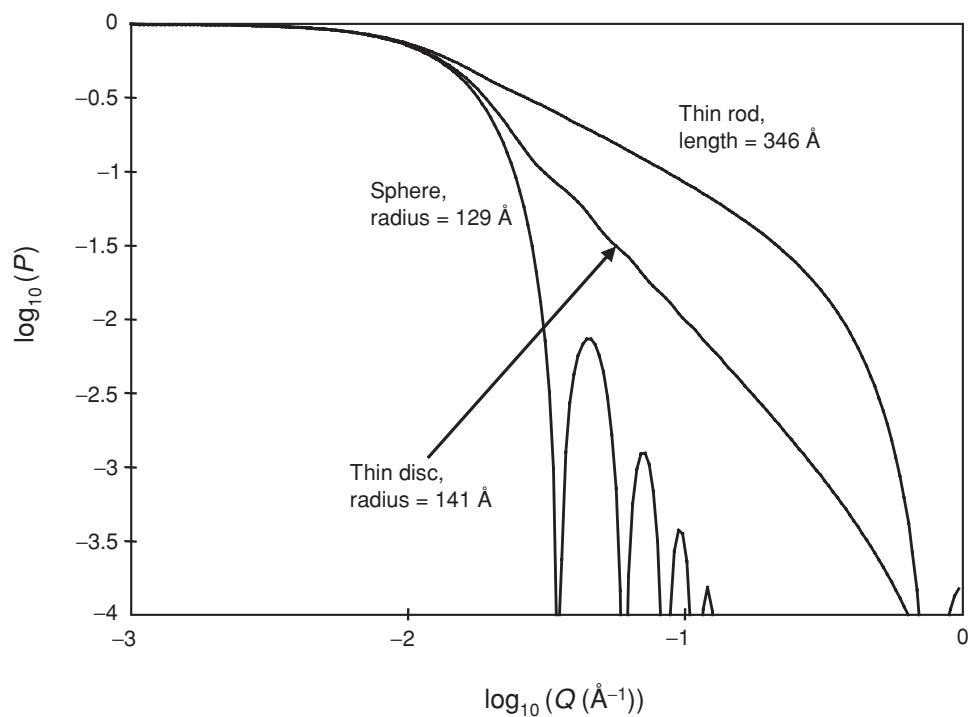
## 12.16 Polydispersity

Polydispersity does not greatly affect the low  $Q$  slope but it tends to smear out the maxima and minima at higher  $Q$  as shown in Figure 12.15. This can be visualised by averaging form factors with slightly different values of the sphere radius,  $R_s$ . The trial and error (fitting) method can be used to deduce the degree of polydispersity.

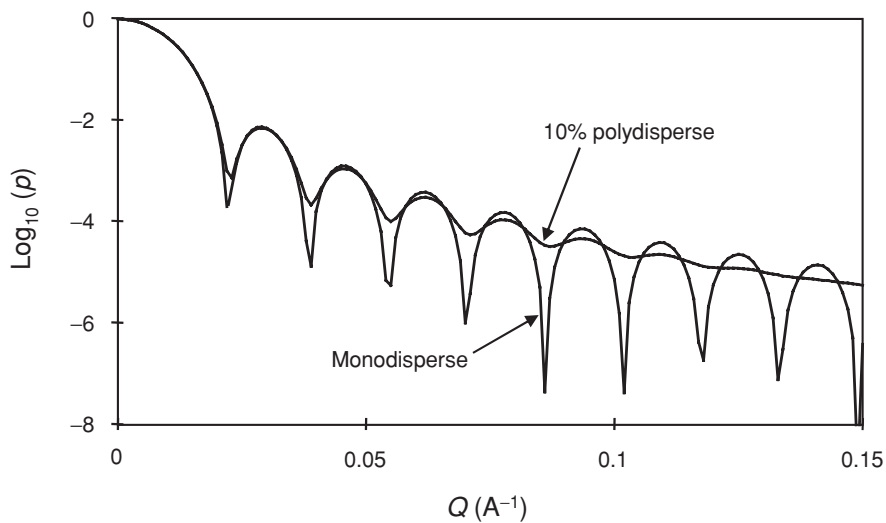
## 12.17 Determination of Particle Size Distribution

There are other computer methods for extracting particle size distributions. For instance there is a maximum entropy approach where the smoothest particle size distribution consistent with the scattering curve is determined [18].

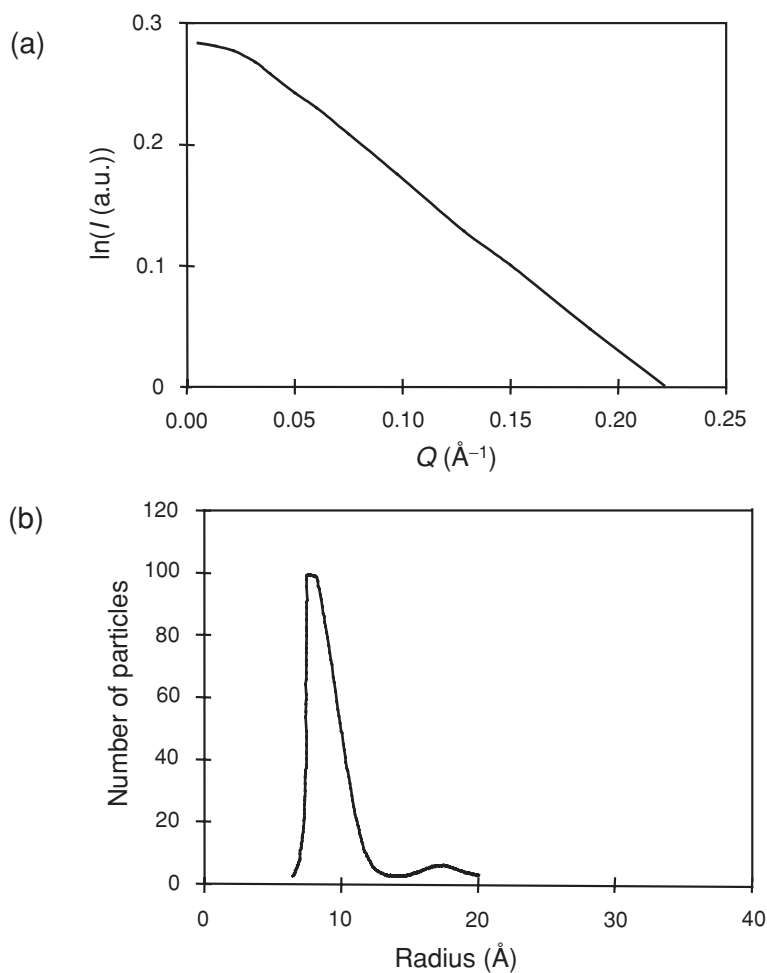




**Figure 12.14** Form factors for different particle shapes with same radius of gyration.



**Figure 12.15** Effect of polydispersity on form factor.

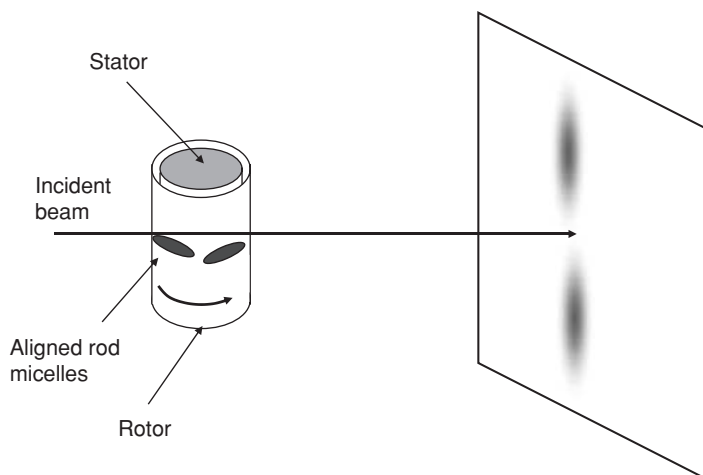


**Figure 12.16** (a) SAXS and (b) particle size distribution determined using maximum entropy method.

Figure 12.16a shows an example of SAXS from partly hydrolysed zirconium chloride which forms polynuclear ions in solution. The maximum entropy particle size distribution is shown in Figure 12.16b. The particles appear to have a radius of 10  $\text{\AA}$  with a small proportion of larger particles (possibly dimers).

## 12.18 Alignment of Anisotropic Particles

For non-spherical particles it is advantageous to align the particles for a small angle scattering measurement. The two characteristic dimensions may then be determined by analysing the scattering in the two perpendicular directions on the detector. For instance, worm-like micelles may be aligned by shearing in a couette as shown schematically in Figure 12.17.



**Figure 12.17** Schematic of scattering from shear aligned sample.

This is usually made of silica which is transparent to neutrons with a gap of 1 mm or less between the external rotor and the internal stator. Nematic liquid crystals may be aligned by applying a magnetic field [19].

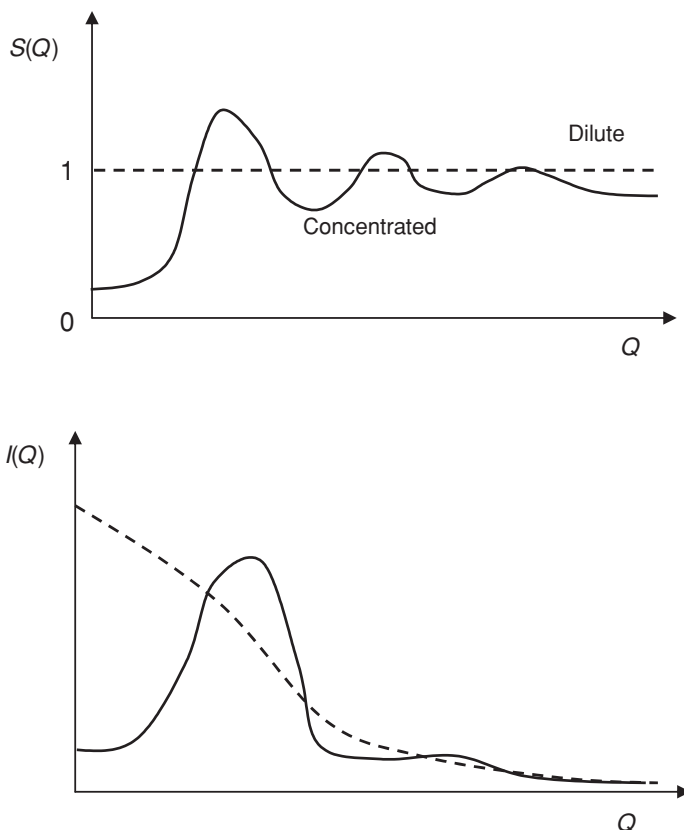
## 12.19 Concentrated Dispersions

For concentrated dispersions, rays scattered from different particles will interfere. This inter-particle interference is accounted for by a term called the structure factor,  $S(Q)$ , as shown in Figure 12.18:

$$I(Q) = (\rho_P - \rho_M)^2 N_P V_P^2 P(Q) S(Q) \quad (12.14)$$

For dilute dispersions,  $S(Q) = 1$ . For concentrated dispersions it is an oscillatory function and it can be used to determine how the particles pack together [20]. Again least-squares model fitting is used to determine parameters such as the closest distance of approach of two particles (the hard sphere repulsion radius). For charged particles (e.g. micelles) the surface charge and screening length may be determined by model fitting [21]. One potential pitfall when using the Guinier law to determine radius of gyration is that the slope of the plot is only  $-R_G^2/3$  if  $S(Q) = 1$ . For samples not very dilute, this may not be correct and analysis using the Guinier law leads to an incorrect value of  $R_G$ .

Figure 12.19 shows SAXS from overbased detergents dispersed in oil. These are used as an engine oil additive. They are calcium carbonate particles stabilised by surfactant. Since the surfactant and the oil have very similar electron densities which are different from that of the  $\text{CaCO}_3$  core, the scattering is dominated by the more electron dense core. For the concentrated dispersion, the peak position and shape may be analysed to give the hard sphere radius. On dilution, the peak disappears ( $S(Q)$  tends to 1) and a Guinier plot can be used to determine the core radius.



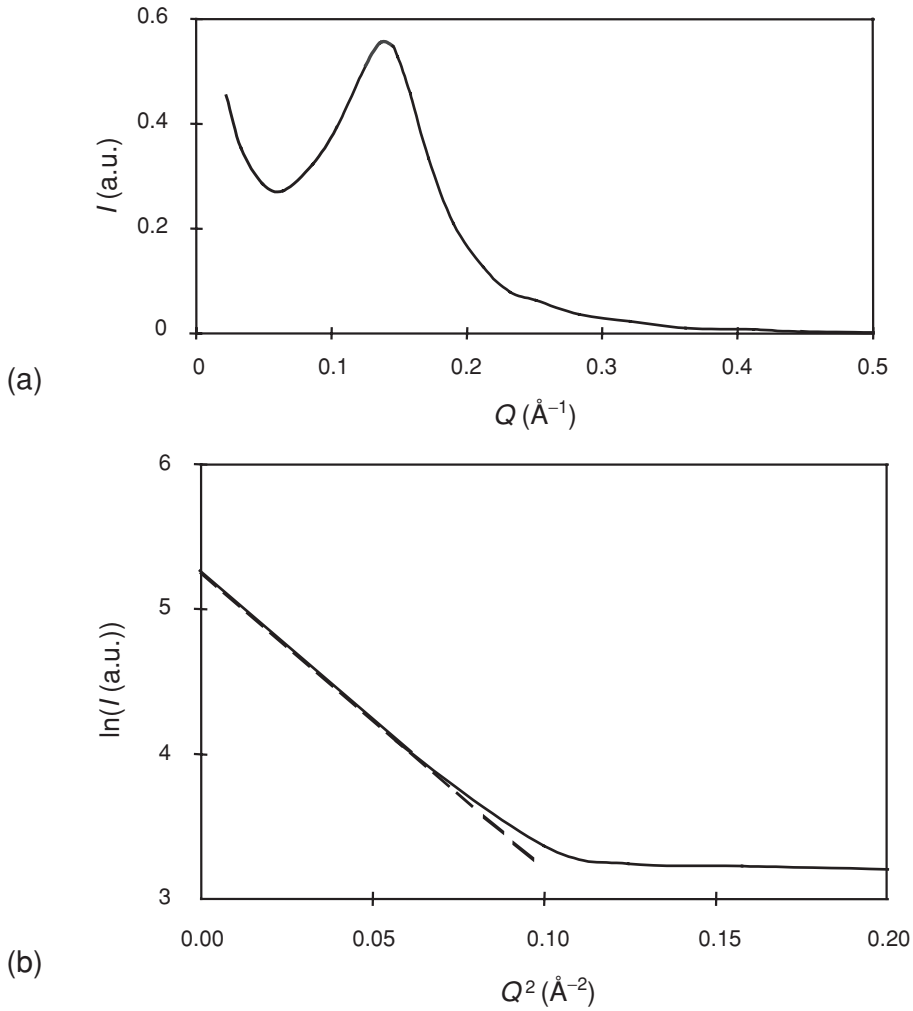
**Figure 12.18**  $S(Q)$  and  $I(Q)$  for dilute and concentrated dispersions.

## 12.20 Contrast Variation using SANS

For SANS the use of contrast variation gives access to more detailed structural information and it is particularly useful for composite particles. Consider such a particle which consists of a core and a relatively thin coating as shown in Figure 12.20. There will be three contributions: scattering from the core, scattering from the coating and scattering that comes from both. This can be modelled but it is complex and there is a tendency for the core scattering to dominate (because core has more volume than coating) so the coating structure is difficult to extract.

To use contrast variation we first arrange for the solvent to have the same scattering length density as the coating. For an aqueous medium, this is done by choosing the correct ratio of  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$ . The coating is now 'contrast matched' and the only scattering is from the core, so radius of gyration of the core,  $R_G$ , can be determined by a Guinier plot.

Now we arrange for the solvent to have the same scattering length density as the core. The core is now 'contrast matched' and the only scattering is from the coating. The thickness of the coating  $R_T$  can now be determined using a version of Guinier's law that applies to the



**Figure 12.19** SAXS from concentrated and dilute calcium carbonate dispersions.

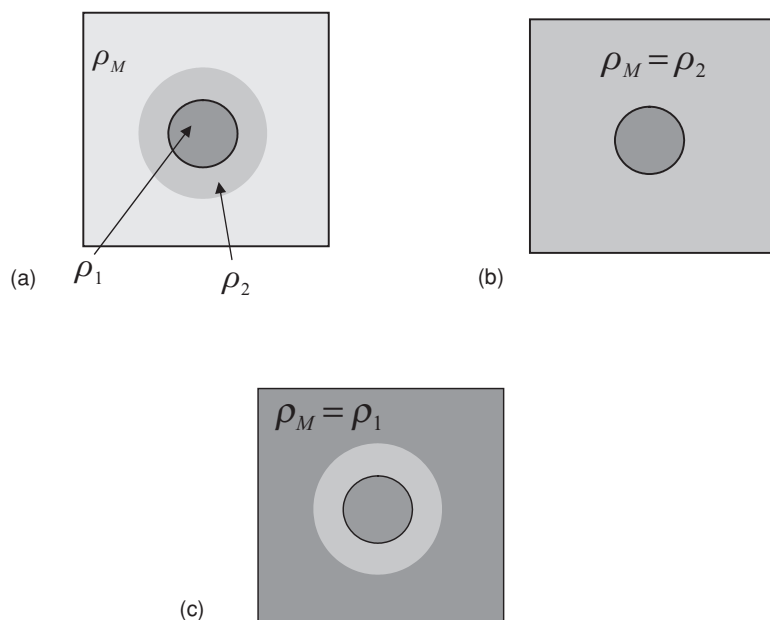
scattering from anisotropic, plate-like objects at  $Q > R_G$  [22]

$$I(Q) \propto \frac{1}{Q^2} \exp(-Q^2 R_T^2) \quad (12.15)$$

so the coating thickness may be determined.

## 12.21 High $Q$ Limit: Porod Law

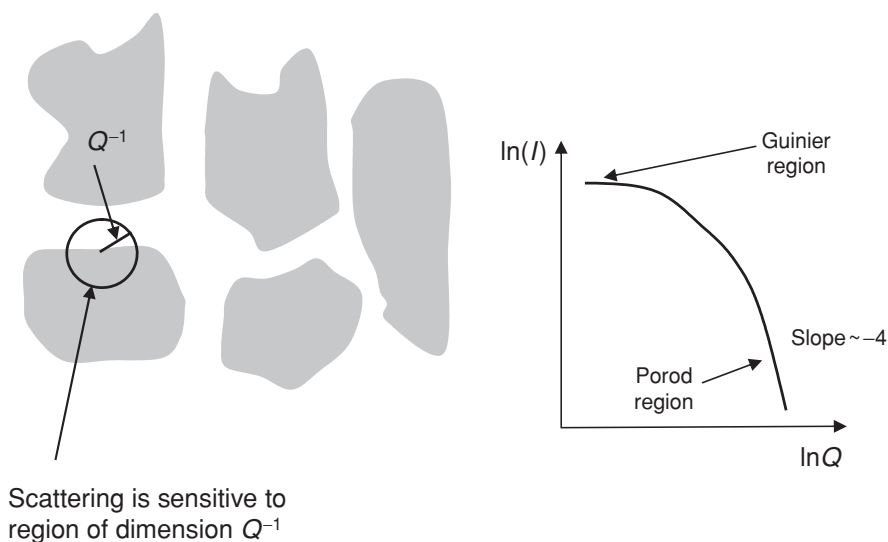
We now consider the form of the scattering at  $Q$  well above the Guinier region. Since the distance probed is inversely proportional to  $Q$ , very high  $Q$  means short distances and the



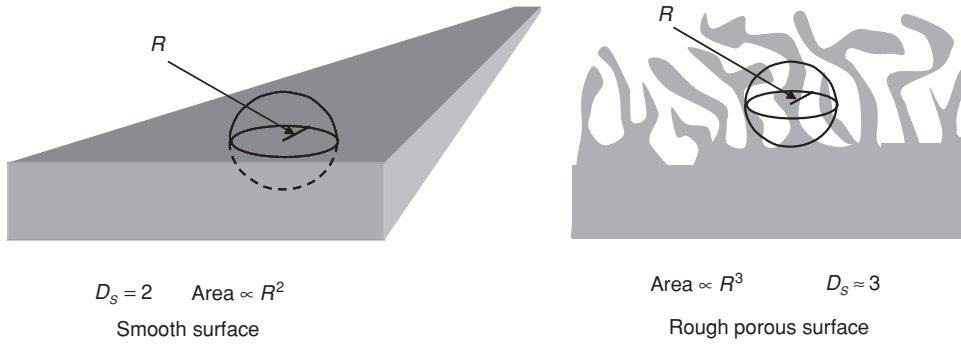
**Figure 12.20** Contrast variations from a composite particle.

only scattering comes from the step in scattering length density at the surface of the particles in a dispersion as shown schematically in Figure 12.21.

There are no inter-particle effects because the distances probed are again very much shorter than the particle separations. This high  $Q$  scattering decays as the fourth power of



**Figure 12.21** Distance probed by high  $Q$  and the corresponding Porod region in the scattering curve.



**Figure 12.22** Fractal surfaces.

$Q$  and its strength depends only on the contrast,  $\Delta\rho^2$ , and the amount of surface area  $S$  in the sample:

$$I(Q) \approx 2\pi S \Delta\rho^2 Q^{-4} \quad (12.16)$$

This is known as Porod's law [23]. The scattering intensity can therefore be used to measure surface area in powders, dispersions, etc.

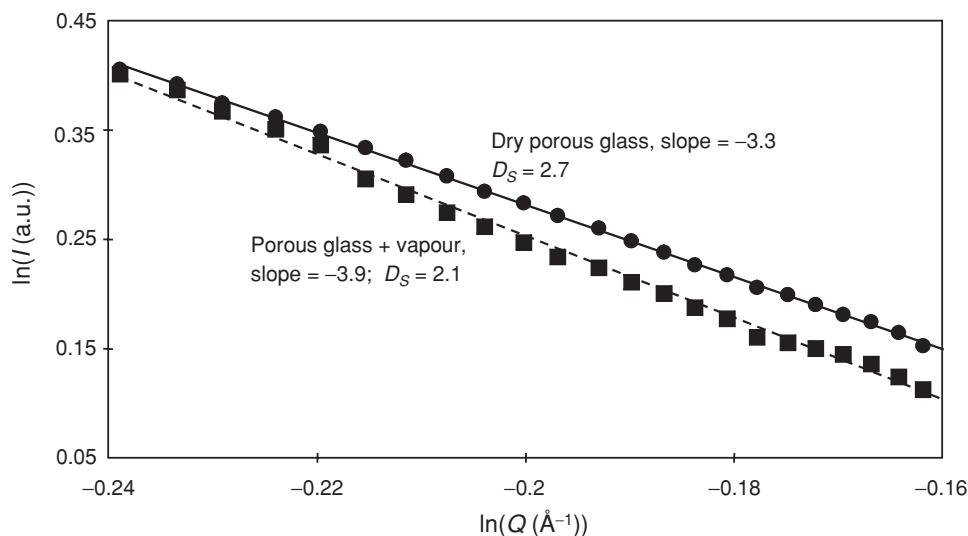
Porod's law is modified if the surfaces are not smooth. The nature of a surface can be characterised by its surface fractal dimension,  $D_s$ . This concept can be understood as follows. Consider a sphere of radius,  $R$ , on a smooth surface. As the radius of the sphere increases, the area of surface in the sphere increases as the second power of the radius. Hence for a smooth surface,  $D_s = 2$ . For a very rough, porous surface the surface area inside the sphere will increase as the third power of the radius. Hence for a rough surface,  $D_s = 3$ . In general the surface fractal dimension will lie between these extremes:  $2 < D_s < 3$ . Figure 12.22 shows the two extremes schematically.

For a fractal surface, Porod's law is extended by changing the power from 4 to  $(6 - D_s)$  and so the fractal dimension may be extracted from the slope of a log-log plot of the high  $Q$  scattering as indicated in Figure 12.22. Note that for a smooth surface, Porod's law is recovered.

$$\ln(I(Q)) \propto -(6 - D_s) \ln(Q) \quad (12.17)$$

Figure 12.23 shows the high  $Q$  scattering from a sample of porous glass (Vycor). When it is dry, the slope of the log-log graph is  $-3.3$ , indicating a surface fractal dimension of  $2.7$  (i.e. quite rough). On exposing it to vapour (a halogenated solvent with similar scattering length density to glass) the slope is  $-3.9$ , indicating a surface fractal dimension of  $2.1$  (i.e. nearly perfectly smooth). The conclusion is that the pores have been filled in by capillary condensed vapour of the same scattering length density as the glass, so the surface appears smooth.

The concept of fractals has many applications. For instance adsorbed polymer layers and aggregates of particles may be characterised by fractal dimensions. A more detailed discussion of scattering from surface and mass fractals may be found elsewhere [24–26].



**Figure 12.23** SAXS from porous glass in dry state and exposed to vapour.

## 12.22 Introduction to X-ray and Neutron Reflection

The reflectivity technique is a recently developed method for studying the structure of macroscopic surfaces [27]. We have seen in Section 12.20 that it is possible to characterise surfaces of particles using small angle scattering by contrast matching the cores of the particles to the solvent. However reflection from macroscopic surfaces has several advantages as compared to studying surfaces of particles in dispersions. These include:

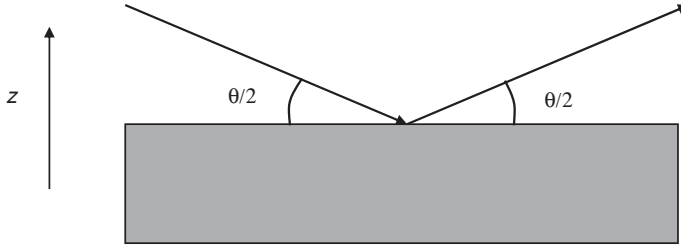
- Reflection is not restricted to stable dispersions.
- Reflection is not restricted to core contrast matched conditions.
- It is more precise because surface contribution to the scattering is separated out experimentally as a specular reflection.
- It is relatively simple to calculate reflectivity from a smooth surface exactly.

The big disadvantage is that at present several square centimetres of flat surface are required. This is simple for the liquid–air interface but more difficult for solid–air, solid–liquid and liquid–liquid interfaces.

## 12.23 Reflection Experiment

The reflectivity method, shown schematically in Figure 12.24, is very simple in principle. A well-collimated monochromatic beam of X-rays or neutrons is brought onto the surface and the intensity of the reflected beam is measured. The angle of incidence is scanned to vary  $Q$ . On a pulsed source, the  $Q$  variation can be done by measuring the time of flight at a fixed angle so that  $Q$  is varied through the range of wavelengths,  $\lambda$ . The reflectivity as a function of  $Q$  is determined and interpreted in terms of the surface structure. It is possible





**Figure 12.24** Principle of reflection experiment.

to purchase an X-ray reflectometer, and X-ray and neutron reflectometers are also available at the central facilities mentioned above.

The reflectivity from a surface  $R(Q)$  may be calculated exactly using a method originally developed for the optics of multi-layers [28]. However, for the purposes of understanding reflectivity results, the kinematic approximation is very useful [29].

In this approximation there are two factors. The first factor is the reflectivity that would be observed from an ideally smooth sharp interface where the change in scattering length density between the two media is  $\Delta\rho$ . It is a  $Q^{-4}$  decay. The second factor results from any surface structure such as an adsorbed layer or diffuseness of the interface. It is the Fourier transform squared of the scattering length density gradient perpendicular to the surface (i.e. the  $z$  direction):

$$R(Q) = \frac{16\pi^2 \Delta\rho^2}{Q^4} \left[ \frac{1}{\Delta\rho} \int_{-\infty}^{\infty} \frac{\partial\rho(z)}{\partial z} e^{iQz} dz \right]^2 = \frac{16\pi^2}{Q^4} \left[ \int_{-\infty}^{\infty} \frac{\partial\rho(z)}{\partial z} e^{iQz} dz \right]^2 \quad (12.18)$$

## 12.24 A Simple Example of a Reflection Measurement

As an example of neutron reflectivity, consider a monolayer of a deuterated surfactant adsorbed at the surface of water as shown schematically in Figure 12.25. The water can be made invisible to neutrons by using 8% by volume  $D_2O$  so that its scattering length density is zero and its contrast matches to air.

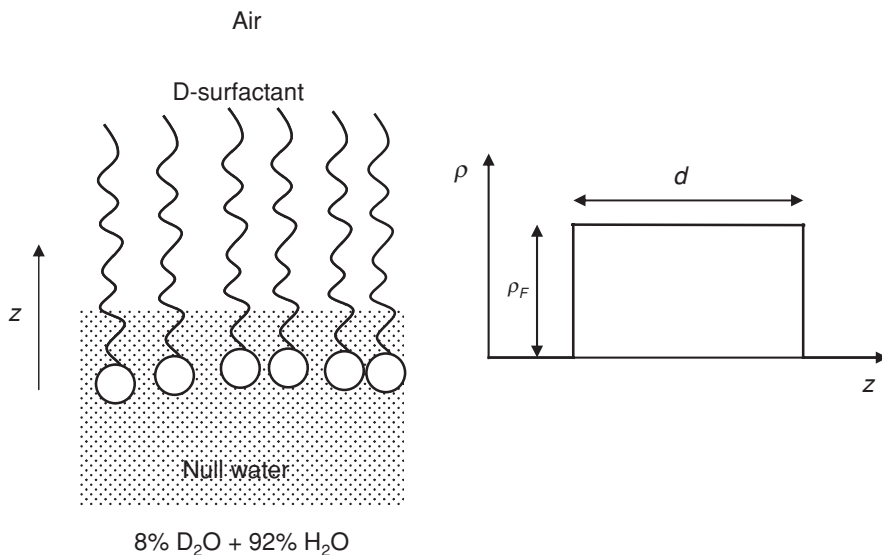
The scattering length density profile is then a simple block shape and it can be shown from standard Fourier transform results [30] that the surface structure factor has a cosine form:

$$R(Q) \approx \frac{16\pi^2}{Q^4} \rho_F^2 2(1 - \cos Qd) \quad (12.19)$$

where  $\rho_F$  is the scattering length density of the surfactant film and  $d$  is its thickness.

If the reflectivity from such a system is plotted as  $RQ^4$ , the rapid decay is removed from the data. The position of the first maximum of the cosine is easily measured and hence the layer thickness,  $d$ , may be determined:

$$d = \pi / Q_{\text{MAX}} \quad (12.20)$$



**Figure 12.25** Deuterated surfactant adsorbed at the surface of ‘null’ water and the corresponding scattering length density profile.

The amplitude of the cosine oscillations is governed by the scattering length density of the film which can therefore be determined:

$$\rho_F = \sqrt{\frac{(RQ^4)_{\text{MAX}}}{64\pi^2}} \quad (12.21)$$

Since the scattering length density depends on the total scattering length of a surfactant molecule,  $\sum_{\text{molecule}} b$ , and the volume it occupies, the area per molecule,  $A$ , may be calculated from the scattering length density of the film:

$$A = \frac{\sum_{\text{molecule}} b}{\rho_F d} \quad (12.22)$$

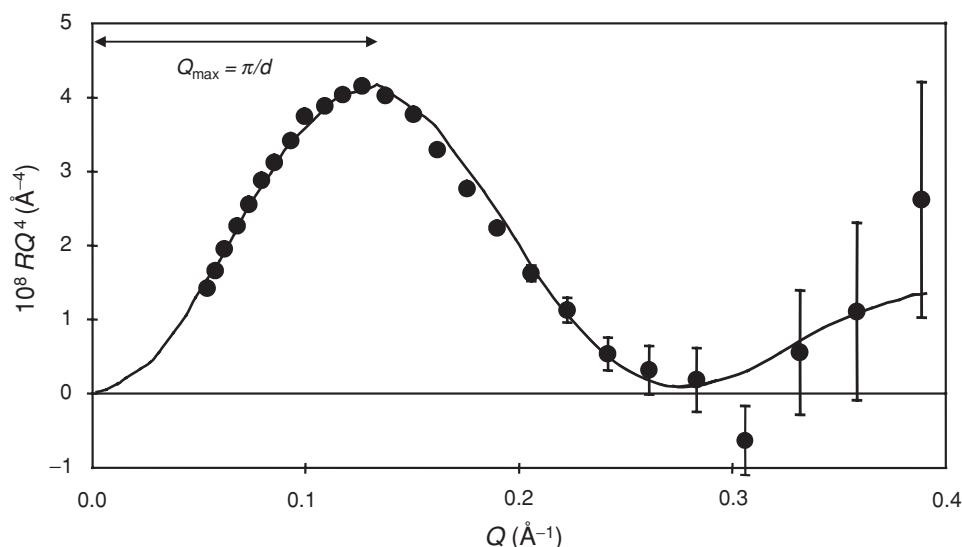
where the summation is over the scattering lengths of all the atoms in one molecule.

Figure 12.26 shows the reflectivity of d-behenic acid spread on water plotted as  $RQ^4$ . The data were taken using the CRISP reflectometer at the ISIS neutron source [31].

The position of the maximum indicates that the layer is  $24 \text{ \AA}$  thick and since the total scattering length of the molecule is  $441 \times 10^{-5} \text{ \AA}^{-1}$ , the area per molecule is determined as  $23 \text{ \AA}^2$ . This simple example shows how the two most important characteristics of the surfactant layer may be determined. The method may be extended to cope with more complex interfaces and to determine more detailed structural information.

## 12.25 Conclusion

We have discussed how light, X-ray and particularly neutron scattering can give useful information on the structure of colloids and surfaces. Although the techniques do not give



**Figure 12.26** Neutron reflectivity of d-behenic acid spread on water.

‘real space’ images, the interpretation of scattering and reflection data in terms of structure is reasonably direct. The data are generally obtained from the samples in an equilibrium state and so are free from artefacts introduced by sample preparation or by the invasive nature of the probe. Hence the methods outlined are widely used in colloid and surface science. Potential users of the methods are encouraged to consult some of the books in the References section where more detailed introductions may be found. The Web pages of the central facilities are also a useful source of further information, particularly about instrumentation.

## References

1. Kostorz, G. (ed.) (1979) ‘Neutron scattering’, *Treatise on Materials Science and Technology*, vol.15. Academic Press, New York.
2. Als-Nielsen, J. and McMorrow, D. (2001) *Elements of Modern X-ray Physics*. Wiley, New York.
3. Kerker, M. (1969) *The Scattering of Light and other Electromagnetic Radiation*. Academic Press, New York.
4. Guinier, A. (1994) *X-ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies*. Dover, New York.
5. Lindner, P. and Zemb, Th. (eds) (2002) *Neutron, X-ray and Light Scattering Methods Applied to Soft Condensed Matter*. North-Holland, Amsterdam.
6. Pusey, P.N. and Taugh, R.J.A. (1982) In R. Pecora (ed.), *Dynamic Light Scattering and Velocimetry: Applications of PCS*. Plenum, New York.
7. <http://www.srs.ac.uk/srs/>
8. <http://www.esrf.fr/>
9. <http://www.diamond.ac.uk/>
10. <http://www.ill.fr/>

11. <http://www.isis.rl.ac.uk/>
12. <http://neutron.neutron-eu.net/>
13. <http://www.ncnr.nist.gov/resources/n-lengths/>
14. Lord Rayleigh, (1911) *Proc. R. Soc. (London)* A, **84**, 25.
15. Guinier, A. (1939) *Ann. Phys.*, **12**, 161.
16. Feigin, L.A. and Svergun, D.I. (1987) *Structure Analysis by Small Angle X-ray and Neutron Scattering*. Plenum, New York.
17. Glatter, O. and Kratky, O. (1982) *Small Angle X-ray Scattering*. Academic Press, New York.
18. Potton, J.A. Daniell, G.J. and Rainford, B.D. (1986) *Inst. Phys. Conf. Ser. No. 81*, Institute of Physics Publishing, Bristol, Chapter 3, p. 81.
19. Hayter, J.B. and Penfold, J. (1984) *J. Phys. Chem.*, **88**, 4589.
20. Ottewill, R.H. (1982) In J.W. Goodwin (ed.), *Colloidal Dispersions*. R.S. Chem. p. 143.
21. Hayter, J.B. and Penfold, J. (1983) *Colloid. Polym. Sci.*, **261**, 1022.
22. Kratky, O. and Porod, G. (1948) *Acta. Phys. Austriaca*, **2**, 133.
23. Porod, G. (1951) *Kolloid-Z.* **124**, 83.
24. Legrand, A.P. (ed.) (1998) *Surface Properties of Silica*. Wiley, New York.
25. Bale, H.D. and Schmidt, P.W. (1984) *Phys. Rev. Lett.*, **53**, 596.
26. Allen, A. and Schofield, P. (1986) *Inst. Phys. Conf. Ser.*, No 81, Institute of Physics Publishing, Bristol, Chapter 3.
27. Bucknall, D.G. (1999) In R.A. Pethick, and J.V. Dawkins (eds), *Modern Techniques for Polymer Characterisation*. Wiley, New York.
28. Penfold, J. and Thomas, R.K. (1990) *J. Phys. Condens. Matter*, **2**, 1369.
29. Als-Nielsen, J. (1885) *Z. Phys. B*, **61**, 411.
30. Champeney, D.C. (1973) *Fourier Transforms and their Physical Applications*. Academic Press, New York.
31. Grundy, M.J., Richardson, R.M., Roser, S.J., Penfold, J. and Ward, R.C. (1988) *Thin Solid Films*, **159**, 43.

# Chapter 13

## Optical Manipulation

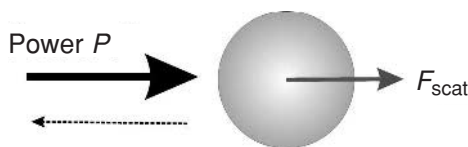
*Paul Bartlett*

### 13.1 Introduction

Manipulating, analysing and organising the mesoscopic structure of materials is probably the most challenging problem currently facing soft matter science. Soft materials, which include polymers, colloids, microemulsions, micellar systems and their aggregates, are characterised by a wide range of length scales, ranging from tens of nanometres to tens of microns, forces from femtonewtons to nanonewtons and timescales which span microseconds to hours. Organising material on these scales has traditionally only been possible through a subtle *chemical* control of interactions and dynamics. However in the last two decades a new generation of optical techniques has emerged that allow soft matter scientists to *physically* reach down into the microscopic world, grab, move and transform dielectric objects at will, with almost nanometre precision. This chapter summarises the ideas behind these new powerful optical manipulation techniques and highlights a few recent applications in soft matter science. More detailed reviews of this area are included in the articles by Grier [1] and Molloy and Padgett [2] while recent developments in optical manipulation techniques have been summarised by Dholakia *et al.* [3].

### 13.2 Manipulating Matter with Light

Moving objects with light seems, at first sight, the stuff of science fiction stories. Indeed ‘tractor beams’ play a major role in classic Sci-Fi tales such as Star Trek where the *U.S.S. Enterprise*’s laser beam is used to pull in crafts, tow another ship or hinder the escape of an enemy spacecraft [4]. The stories may be a little fantastical but the science is sound. Light can move matter because photons carry a momentum. Each photon of wavelength  $\lambda$  has a momentum  $p = h/\lambda$  where  $h$  is Planck’s constant. Illuminating an object leads to a change in the direction of light as a result of refraction, reflection and diffraction. The incident momentum of the beam of photons is changed and so from Newton’s laws of motion the object must also experience a force. Of course, the forces are not large enough to move spaceships (unless the beam is phenomenally intense) but for small objects, such as micron-sized particles, the forces are sufficient to allow them to be moved at will.



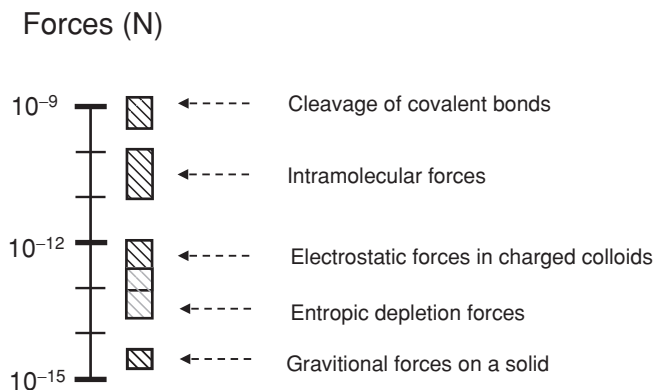
**Figure 13.1** Optical scattering force on a sphere.

A very simple calculation, illustrated in Figure 13.1, confirms the magnitude of optical forces. Imagine a light beam of power  $P$  incident on a microscopic sphere. As every photon carries an energy  $h\nu$  the number of photons incident on the particle per second is  $P/h\nu$ . If a fraction  $q$  of the beam is reflected back then the momentum transferred to the particle leads to a scattering force on the particle of magnitude

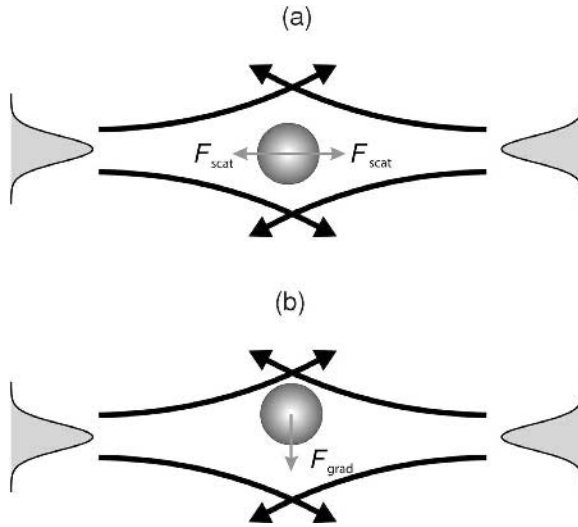
$$F_{\text{scat}} = 2 \left( \frac{P}{h\nu} \right) \times \left( \frac{h}{\lambda} \right) \times q = \frac{2qnP}{c} \quad (13.1)$$

where  $c$  is the velocity of light and  $n$  is the refractive index of the medium. Inserting typical values gives a crude estimate of the strength of optical forces. Assuming, for instance, a 100 mW laser beam focused on a dielectric sphere of radius  $\lambda$ , for which  $q$  is of order 0.05 [5], yields an optical force of about 40 pN. Although this force is obviously far too small to move anything as large as a spacecraft, at the microscopic level such piconewton-level forces can have a very significant effect. To see this we need to consider the typical magnitudes of forces found in colloidal systems.

Figure 13.2 shows schematically the range of typical forces encountered in soft matter science. Rupture of covalent bonds requires forces of order 1–2 nN while forces of about 20–50 pN are sufficient to unravel polymer chains, convert DNA from a double helix to a ladder or to break most van der Waals interactions. Colloidal forces are typically an order of magnitude smaller with the interaction forces between micron-sized colloidal particles being on a scale of a few ten to hundreds of femtonewtons. Probably the weakest forces encountered are those originating from gravity, where the sedimentation force on a colloidal particle is



**Figure 13.2** The strengths of forces encountered in soft matter science.

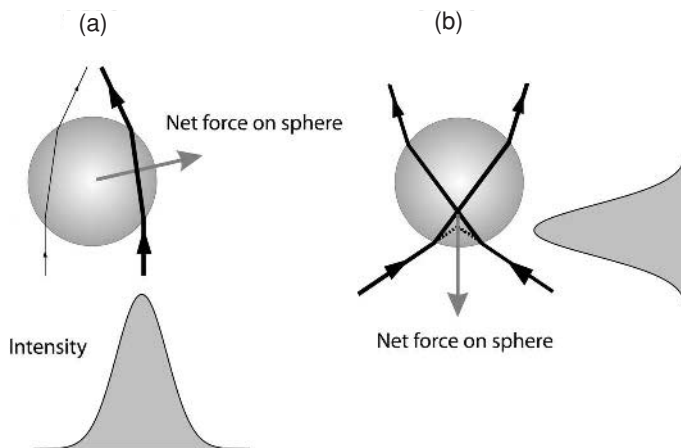


**Figure 13.3** An optical trap generated by two counter-propagating beams. (a) At the equilibrium position the axial scattering forces  $F_{\text{scat}}$  generated by each beam exactly balance, while (b) when the trapped particle is displaced radially there is an unbalanced lateral gradient force  $F_{\text{grad}}$ .

typically a few femtonewtons. Optical techniques, as we shall see below, may be used to apply and measure forces ranging between  $\sim 10$  fN and  $\sim 100$  pN making them ideally suited to the study and manipulation of soft matter.

The first three-dimensional optical traps were built by Arthur Ashkin working at Bell Labs in the early 1970s [6]. The trap consisted of two counter-propagating weakly diverging laser beams as shown schematically in Figure 13.3. At the equilibrium point the axial forces  $F_{\text{scat}}$ , generated by the *scattering* of photons from both beams, exactly balance and the particle is stably trapped. Any motion along the axis leads to a net scattering force which moves the particle towards the equilibrium point (Figure 13.3a). Although this might be expected from our discussion above, Ashkin saw, rather surprisingly, that the particle was also confined radially (Figure 13.3b). This observation, although apparently mundane, is key to the successful development of optical trapping techniques. It was the first demonstration that radiation pressure could also produce a *transverse* force component, which acts perpendicular to the line of the beam. The transverse or *gradient* force,  $F_{\text{grad}}$ , acts to move the particle to wherever the laser field is the highest. So in the case of the counter-propagating trap any radial displacement of the particle is opposed by gradient forces generated by both beams. The particle is stable against random displacements in all three dimensions and ‘optical trapping’ had arrived in the laboratory.

While the counter-propagating trap worked, alignment was rather tedious and the need to get optical access from two sides restricted its use. It was therefore a significant breakthrough when Ashkin showed in 1986 [7] that the gradient forces produced near the focus of a *single* tightly focused laser beam could trap a transparent particle in three dimensions. In the last two decades, this technique, now referred to almost universally as ‘optical tweezers’, has become a mainstream tool in nanotechnology and biology.



**Figure 13.4** Force generation in a single-beam optical gradient trap.

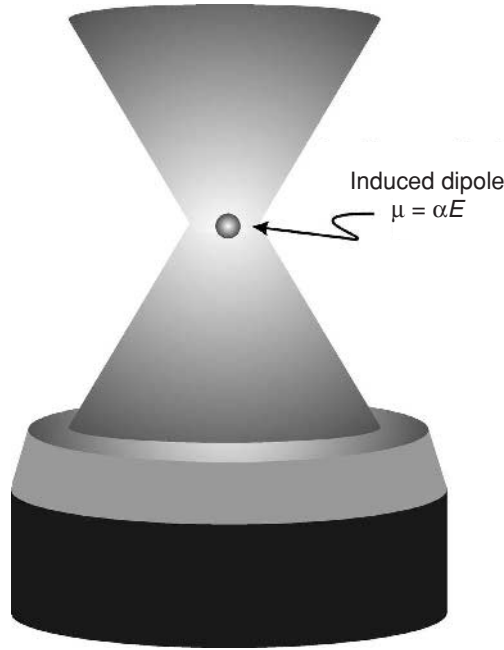
### 13.3 Force Generation in Optical Tweezers

To understand how a single tightly focused laser trap works, study Figure 13.4. This figure shows the passage of light rays as they travel through a transparent sphere with a high index of refraction. At the surface of the sphere light is refracted and is bent towards the normal on entry and away from the normal on exit, according to Snell's law. Each time the light ray is refracted there is a change in photon momentum and so from Newton's third law the particle will experience an equal and opposite force. Figure 13.4a illustrates the case when the particle is displaced to the left of the centre of the beam. More light is refracted to the left than to the right so the particle experiences a net force directed to the right and towards the centre of the beam. When the particle is moved up, away from the beam focus, as shown in Figure 13.4b, then light rays are refracted upwards which generates a reaction force on the sphere which pulls it towards the beam focus. The net effect is that the motion of the particle is constrained in all three dimensions.

This ray optics approach gives remarkably accurate estimates for the strength of optical trapping provided the sphere is significantly larger than the wavelength of the laser [8]. For smaller particles it is better to use arguments based on the strength of the electric field at the trapped particle [9]. Focusing a laser beam generates an intense electric field at the beam focus, as shown schematically in Figure 13.5. The effect of the electric field is to polarise the trapped sphere and generate a time-dependent induced dipole  $\mu = \alpha E$ , whose size depends upon the polarisability  $\alpha$  of the sphere. To first order, the polarisability  $\alpha$  of a sphere of radius  $r$  varies like  $\alpha \approx (n_p - n_m)r^3$ , where  $n_p$  and  $n_m$  are the index of refraction of the particle and medium respectively. So a sphere with an index of refraction above that of the medium has a positive polarisability.

The effect of the oscillating induced dipole is twofold. First it emits radiation which gives rise to a scattering force on the particle. The intensity of light scattered is proportional to the square of the induced dipole, so the strength of the scattering force scales as the square





**Figure 13.5** A sphere trapped by the intense electric field at the focus of a single-beam gradient trap.

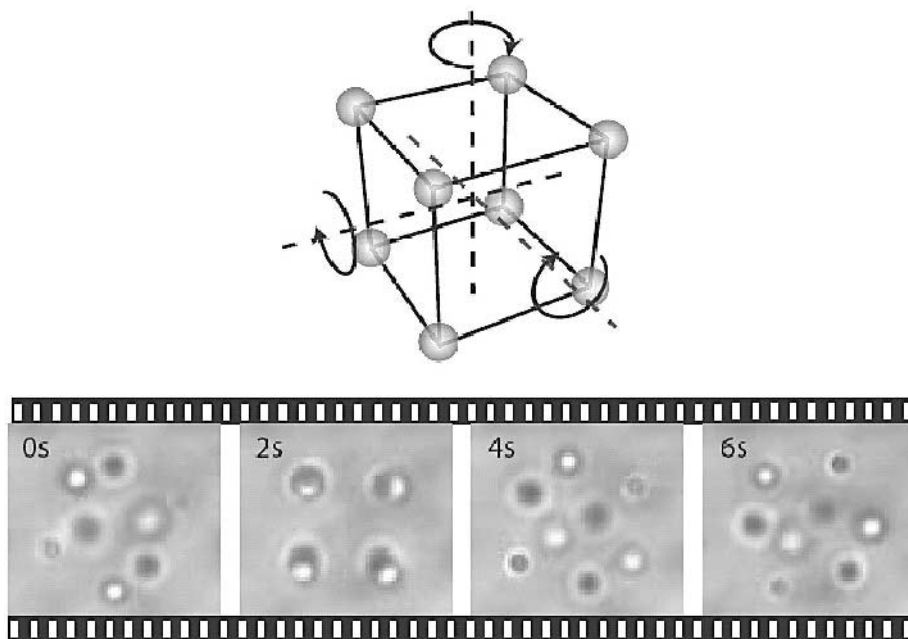
of the polarisability,  $F_{\text{scat}} \approx \alpha^2 E^2 \approx (n_p - n_m)^2 r^6 I$  where  $I$  is the laser intensity. Although not apparent from this equation,  $F_{\text{scat}}$  is parallel to the direction of light propagation so the particle is guided along the beam by the scattering force. The second effect arises from the field gradient present near the beam focus. An aligned dipole, which consists of a separated positive and negative centre of charge, experiences no net force when placed in a uniform electric field. However in a spatially varying field the electric fields at the positive and negative centres of charge differ so that there is a net force on the dipole. The strength of this gradient force is  $F_{\text{grad}} \approx \mu \nabla E$  which using our earlier expression for the induced dipole is easily seen as equivalent to  $F_{\text{grad}} \approx (\alpha/2) \nabla E^2$  ( $\nabla$  is a differential operator). The gradient force is linear in the polarisability and is directed towards the region of highest light intensity in the case where  $\alpha > 0$ . For stable three-dimensional trapping we must concentrate on increasing the gradient forces so that they exceed the scattering forces. This cannot be done by simply adjusting the laser intensity since both the scattering and gradient forces vary linearly with intensity. Instead we need to maximise the intensity gradient near the beam focus. This is most readily achieved by using a high numerical aperture microscope objective which brings light from a wide cone angle to a sharp focus. Finally we emphasise that the refractive index of the trapped particle must be higher than that of the surrounding medium to ensure that the gradient force is directed towards the maximum intensity region. An air bubble or generally a particle with a low refractive index is expelled from the beam focus as the gradient forces are reversed.

## 13.4 Nanofabrication

Optical tweezers offer a highly controlled way of manipulating soft matter systems. There is no direct physical contact with the system, so there is no possibility of contamination. Furthermore, the ability to remotely position colloids in space means that it is now possible to construct new classes of materials. One of the most exciting developments in optical tweezer technology has been the creation of three-dimensional arrays of optical tweezers. These multiple trap systems are created by using a computer-controlled liquid crystal spatial light modulator to generate a highly controlled phase modulator. When illuminated with a single coherent laser beam the outgoing reflected beam is precisely modulated in phase so that when focused in the tweezer plane there is constructive interference between different parts of the beam and an array of tweezers is generated. Using this technique (known as holographic optical tweezers or HOT) more than nearly 2000 traps have been generated in a plane [10].

Holographic optical tweezers are computer controlled and can be reconfigured rapidly so that the array of traps can be adjusted in space and the structure rotated or modified in real time. Figure 13.6 shows, for instance, a sequence of video images of eight spheres trapped at positions corresponding to the corners of a 'tumbling' cube where the resolution of the spatial light modulator allows the unit cell size to be set arbitrarily between 4 and 20  $\mu\text{m}$ .

Such controllable three-dimensional patterning is currently under active investigation as a means of organising nanoparticles into photonic band-gap materials.



**Figure 13.6** Eight silica spheres (diameter 2  $\mu\text{m}$ ) trapped at the corners of tumbling simple cubic lattice [11] (reproduced with permission).

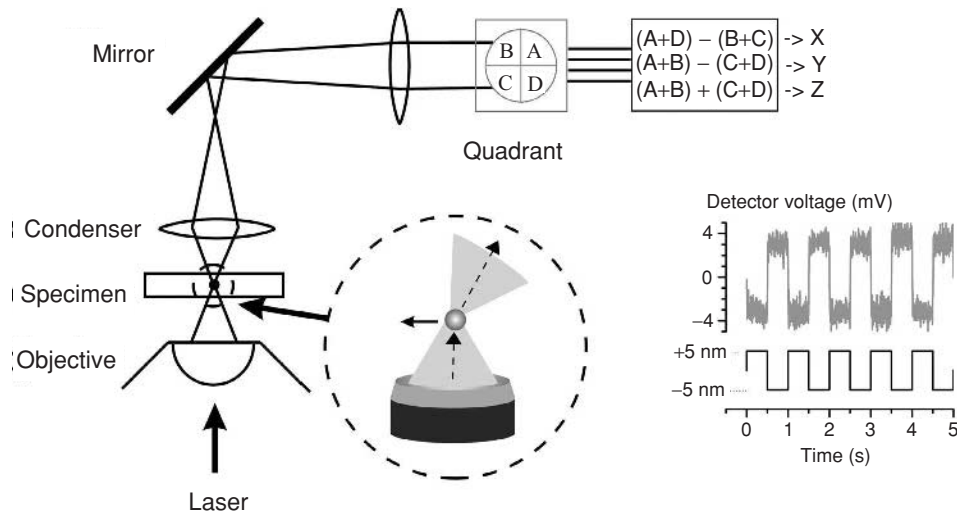
## 13.5 Single Particle Dynamics

For the last few decades much of what we know about the dynamics of soft matter has been derived from bulk techniques such as light, x-ray and neutron scattering which measure the properties of a statistically very large ensemble of particles (typically  $> 10^{12}$ ). The averaging inherent in these measurements means that one cannot measure the entire distribution of particle properties. Single particle techniques offer a wealth of new data on individual properties which allow much more stringent testing of ideas and also offer the potential to reveal entirely new behaviour that is not discernible in averaged results particularly from heterogeneous populations. In this section we show how optical tweezer techniques allow us to measure the dynamics of individual colloidal particles with nanometre spatial and microsecond temporal resolution. We illustrate the power of this method by exploring the heterogeneous dynamics of colloidal gels.

### 13.5.1 Measuring nanometre displacements

The three-dimensional position of a sphere held in an optical trap can be measured with a resolution of a few nanometres using the four-quadrant photosensor depicted in Figure 13.7.

The sensor relies on interference between the light scattered forward by the particle and the transmitted trapping laser [12]. Motion of the particle within the trap changes the direction of the scattered light and so alters the interference pattern. The interference image is projected onto a four-sector quadrant photodetector. The resulting photocurrents are amplified, and combined to yield voltage signals which are proportional to the X, Y and Z coordinates of the trapped sphere. Because of the intense illumination the resolution is very high and using low noise electronics it is possible to achieve nanometre resolution over



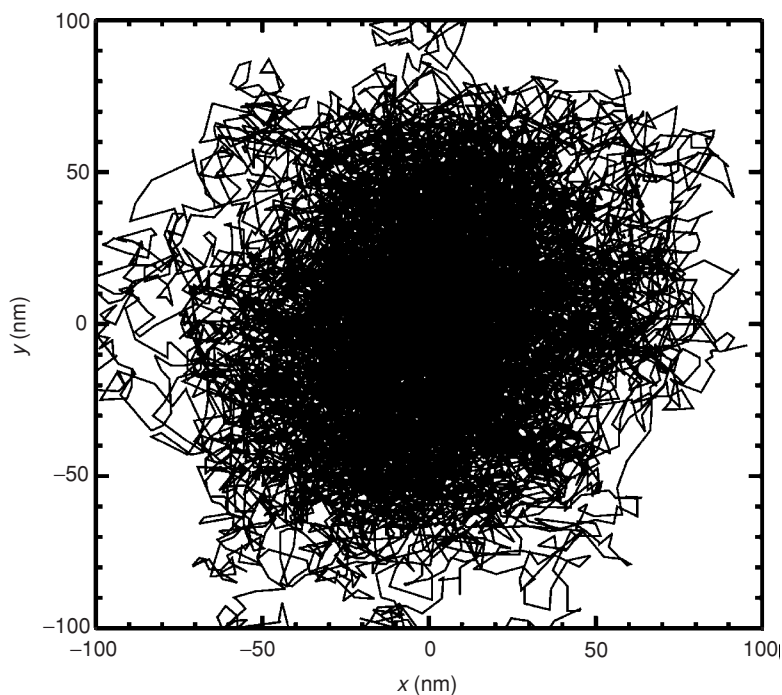
**Figure 13.7** Quadrant photosensor detection of particle displacement.

a bandwidth from about 1 Hz to 10 kHz. Figure 13.7 shows the detector response when a sphere is scanned back and forth in 10 nm steps. Clearly the noise level is at the nanometre level.

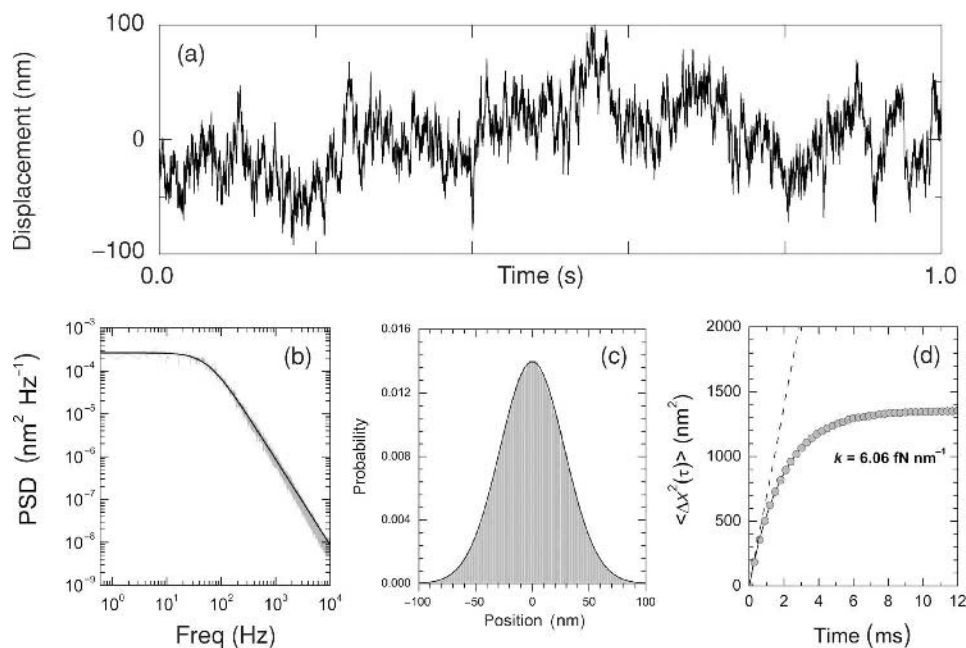
### 13.5.2 Brownian fluctuations in an optical trap

A particle held within an optical trap is in fact not fixed but fluctuates in position as a result of a balance between thermal Brownian forces and the optical gradient and scattering forces. The quadrant photosensor provides a very accurate picture of these thermal fluctuations. Figure 13.8 shows an example of the chaotic random trajectory measured for a particle held within an optical trap which can fluctuate about 50 nm around the optical axis.

Analysis of the time-dependent fluctuations of the trapped particle provides a quick and accurate method to characterise the strength of the optical trap. The gradient force provides a restoring force which, over distances of several hundred nanometres, is a linear function of the displacement  $x$  of the particle from the centre of the beam focus so that  $F_{\text{grad}} = -kx$  where  $k$  is the stiffness or force constant of the optical trap. The trapped particle is essentially bound by a weak spring to the centre of the trap. The particle however does not oscillate (as it would in air or a vacuum) because the motion is heavily damped by the surrounding viscous liquid medium. Figure 13.9a shows that the particle fluctuates



**Figure 13.8** The Brownian trajectory (measured for 0.8 s) of a PMMA microsphere of diameter  $0.8\ \mu\text{m}$  held in an optical trap.



**Figure 13.9** The thermal fluctuations of a particle held within an optical trap.

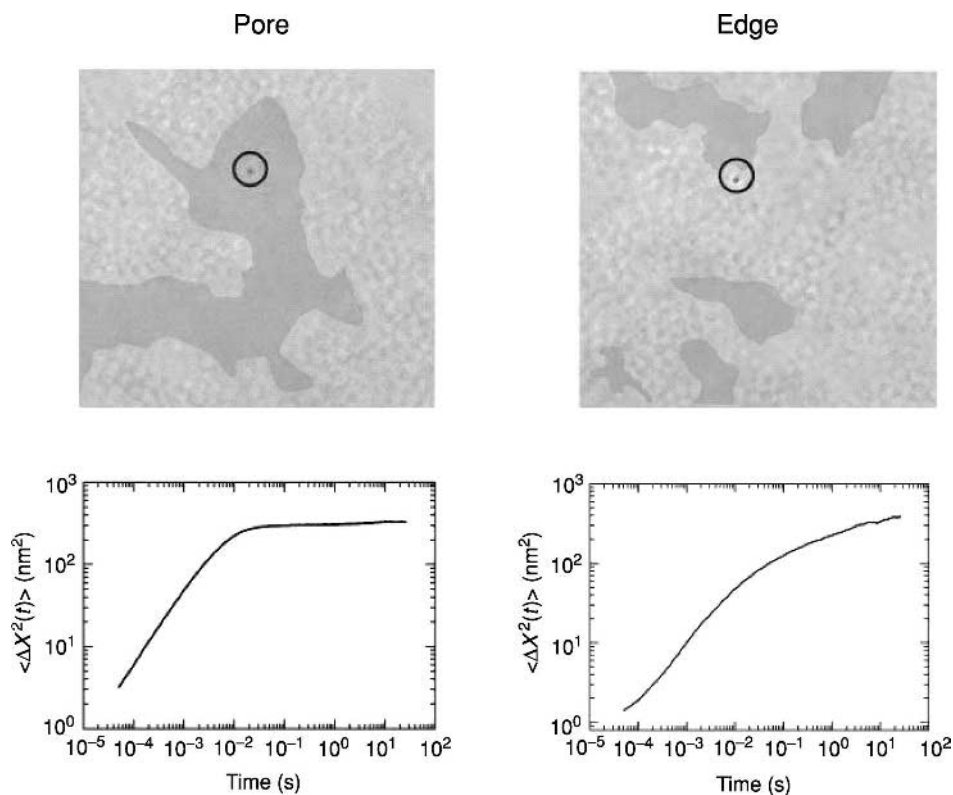
on a timescale of order 10 ms but the motion is very erratic as a consequence of random solvent collisions. The statistics of the thermal fluctuations however reveal the nature of the interaction between the particle and the optical trap. Spectral analysis (which reveals the strength of each Fourier component) of the motion (shown in Figure 13.9b) shows a flat plateau at low frequencies with a high frequency ( $\omega$ ) Lorentzian decay (proportional to  $\omega^{-2}$ ), characteristics of thermal fluctuations in a harmonic potential. The solid line shows that the measured fluctuations are well described by theory. Similar information is obtained from the probability of finding the particle a distance  $x$  away from the beam centre, which is readily calculated from the measured trajectory, an example of which is shown in Figure 13.9c. This distribution yields directly the optical trap stiffness  $k$  since from the Boltzmann law  $P(x) = A \exp(-kx^2/2k_B T)$ . Finally Figure 13.9d depicts the mean-square displacement (MSD),  $\langle \Delta x^2(t) \rangle = \langle |x(t) - x(0)|^2 \rangle$ , of the particle as a function of time. Fitting the measured MSD gives the stiffness of the optical trap as  $6 \text{ fN nm}^{-1}$ . Consequently a typical position resolution of 1 nm equates to a force resolution of about 6 fN.

### 13.5.3 Dynamical complexity in colloidal gels

Particulate gels, produced by adding a non-adsorbing polymer to a stable colloidal suspension, play a wide role in industry. Gels are intrinsically complex, soft, multi-phase systems with an internal organisation or microstructure which varies with length scale. It is this structural complexity which has frustrated previous attempts to link the bulk properties of

gels to what is happening on the individual particle scale. An understanding of this local environment is critical for processing and understanding the long term stability of these materials. Particles with short-ranged attractions show an abrupt change in dynamics as the volume fraction is increased. At the arrest transition, the system transforms from a viscous liquid to a jammed, structurally disordered solid capable of sustaining a shear stress. While many commonplace materials such as foods, pesticides, coatings and cosmetics consist of colloid or protein gels, the molecular mechanism of arrest is a subject of intense scientific debate. The scientific challenge is to explain the rich variety of arrested states seen in protein and colloid gels. Optical tweezers provide a unique method to explore the very different microenvironments present within a gel.

Figure 13.10 shows the trajectory of a single spherical titania particle trapped within a gel, formed from equal sized ( $1.3\ \mu\text{m}$ ), density and index-matched poly(methyl methacrylate) spheres. The different microenvironments present within a gel sample are characterised by positioning the probe particle within a pore and at the edge of the particle chains. As is clear from the figure the different microenvironments result in very different particle dynamics. A scattering measurement, for instance, would average these different dynamics together and not reveal the true heterogeneous nature of the system.



**Figure 13.10** The mean-square displacement of a single colloidal particle located within the pore space and at the edge of a particle chain. Note the significant differences between the two microenvironments.

## 13.6 Conclusions

This chapter has provided a short introduction to the physics of optical tweezers and given a few simple applications of the technique to soft matter science. Optical tweezers can now trap and orient a large number of particles and measure their properties with high precision. The optical toolkit is now in place. In the next few years, we can expect to see these techniques rapidly developing to the point where they become mainstream, providing researchers with the ability to control the microscopic world with unparalleled precision.

## References

1. Grier, D.G. (2003) A revolution in optical manipulation. *Nature*, **424** (6950), 810–816.
2. Molloy, J.E. and Padgett, M.J. (2002) Lights, action: optical tweezers. *Contemp. Phys.*, **43** (4), 241–258.
3. Dholakia, K, Spalding, G.C. and MacDonald, M.P. (2002) Optical tweezers: the next generation. *Phys. World*, **15** (10), 31–35.
4. Krauss, L.M. (1997) *The Physics of Star Trek*. Harper Collins, London.
5. van de Hulst, H.C. (1981) *Light Scattering by Small Particles*. Dover, New York.
6. Ashkin, A. (2000) History of optical trapping and manipulation of small-neutral particles, atoms, and molecules. *IEEE J. Sel. Top. Quantum Electron.*, **6** (6), 841–856.
7. Ashkin, A., Dziedzic, J.M., Bjorkholm, J.E. and Chu S. (1986) Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.*, **11**, 288–290.
8. Ashkin, A. (1992) Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime. *Biophys. J.*, **61** (2), 569–582.
9. Harada, Y. and Asakura, T. (1996) Radiation forces on a dielectric sphere in the Rayleigh scattering regime. *Opt. Commun.*, **124**, 529–541.
10. Curtis, J.E., Koss, B.A. and Grier, D.G. (2002) Dynamic holographic optical tweezers. *Opt. Commun.*, **207** (1–6), 169–175.
11. Leach, J., Sinclair, G., Jordan, P., Courtial, J., Padgett, M.J., Cooper, J., *et al.* (2004) 3D manipulation of particles into crystal structures using holographic optical tweezers. *Opt. Express*, **12** (1), 220–226.
12. Gittes, F. and Schmidt, C.F. (1998) Interference model for back-focal-plane displacement detection in optical tweezers. *Opt. Lett.*, **23** (1), 7–9.

# Chapter 14

## Electron Microscopy

*Sean Davis*

### 14.1 General Features of (Electron) Optical Imaging Systems

The aim of any imaging system is to produce an image from an object. Generally this also involves magnifying the image, but the main limitation on the maximum useful magnification is the resolution, which is defined as ‘the smallest distance between two adjacent points which can be seen as separate’.

The human eye can distinguish features separated by distances of  $\sim 0.2$  mm. The resolving power of a light microscope is diffraction limited and given by

$$r = 0.61\lambda/\mu \sin \alpha \quad (14.1)$$

where  $\lambda$  is the wavelength of radiation,  $\mu$  is the refractive index and  $\alpha$  is the semi-angle subtended at the specimen.

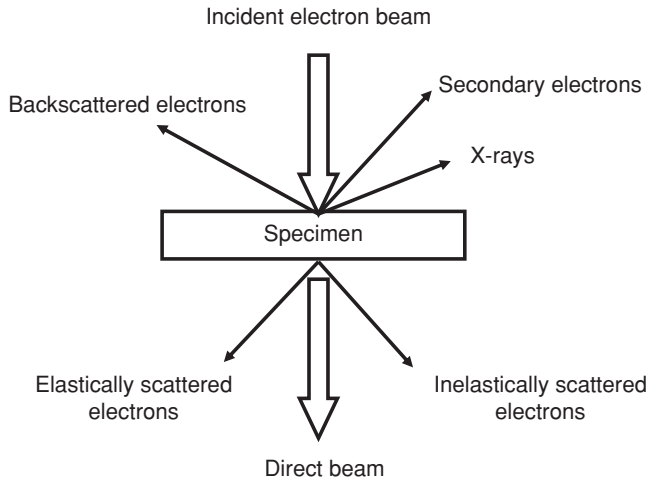
Generally this corresponds to a resolution of about 200 nm, i.e. magnifications of around  $1000\times$ . Electron microscopy offers potentially improved resolution primarily due to the shorter wavelength of electrons compared to light. When electrons interact with a specimen a number of different signals can be produced, the most common of which are shown in Figure 14.1.

Historically, the first electron microscopes to be developed were transmission electron microscopes (TEM) in 1930, these electron optical systems being conceptually very similar to light microscopes (see Table 14.1).

As well as imaging, TEM is used to obtain crystallographic information about samples using electron diffraction. The first scanning electron microscope (SEM) was not produced until 1960. The two main types of signal used for imaging in SEM are the low energy secondary electrons (SE) and elastically scattered backscattered electrons (BSE) (see Figure 14.2). Although the mechanism of image formation is very different to TEM, the images are easy to interpret, and the machines share a number of common components (electron gun, vacuum system, electromagnetic lenses).

The design of conventional electron microscopes imposes a number of restrictions on the nature of the specimens that can be imaged. The most obvious restriction is that in conventional electron microscopy (SEM or TEM) the specimens have to be vacuum stable (at room temperature and pressure the path length of electrons would only be a few mm). In addition the samples have to be stable to the electron beam (i.e. not thermally or photo-sensitive). These restrictions can be particularly important when imaging colloidal systems, and will be discussed further in the following sections.

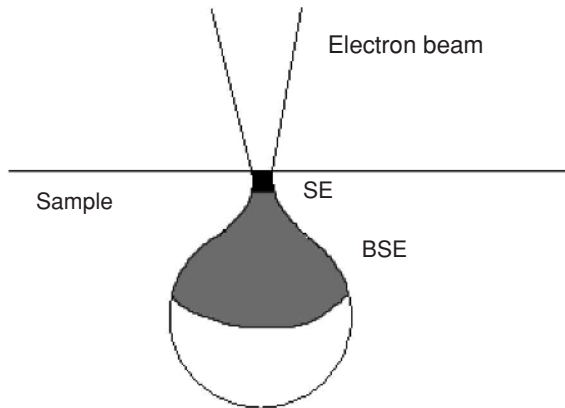




**Figure 14.1** Specimen electron interactions.

**Table 14.1** A comparison of optical and electron microscopes

<b>Illumination</b>	Electron beam	Light
<b>Wavelength</b>	0.0086 nm (20 keV) 0.0025 nm (200 keV)	750 nm visible
<b>Construction</b>	Vacuum	Atmosphere
<b>Lens type</b>	Electromagnetic	Glass
<b>Aperture angle</b>	35 min	70°
<b>Resolving power</b>	0.2 nm	~200 nm
<b>Magnification</b>	$10^{-1} \times 10^6$ ; variable	10–200; fixed lens
<b>Focusing</b>	Electronic	Mechanical
<b>Contrast</b>	Scattering, absorption, diffraction, phase	Absorption reflection, phase, polarisation



**Figure 14.2** Volume of interaction of electron beam with sample.

## 14.2 Conventional TEM

### 14.2.1 Background

In TEM the electron beam is passed through a thin specimen, and the transmitted image is magnified and focused. Because electromagnetic lenses behave as thin lenses, electron optics can be treated similarly to light optics. However, for EM lenses,  $\mu = 1$ , and the electrons are deflected through very small angles so  $\sin \alpha = \alpha$ . Therefore,  $r = 0.61\lambda/\alpha$ .

Resolution increases as wavelength decreases (increased accelerating voltages), and also as the objective lens aperture is made larger. For an instrument operated at an accelerating voltage of 50 keV ( $\lambda = 0.0055$  nm) and  $\alpha = 0.1$  rad the theoretical resolution is  $r = 0.03$  nm, which is subatomic.

In fact the actual resolution obtainable is typically of the order of 0.2–0.3 nm. The theoretical resolution is never obtained due to lens aberrations which are difficult to fully compensate for in electron optical systems. The two main sources which limit the resolution are chromatic aberrations and achromatic (spherical) aberrations. Both of these types of aberrations result in the electron beam being focused in a range of positions along the ‘optical’ axis due to differences in the energy of the electrons (chromatic) or path length travelled (achromatic). Although chromatic aberration can be reduced using monochromatic sources, inelastic scattering means the electrons emerging from the specimen have a spread of energy. Thinner specimens and higher accelerating voltages reduce the number of inelastic scattering events, and enable higher resolution imaging. Achromatic aberrations can be reduced by using small objective apertures, i.e. selecting electrons close to the ‘optical’ axis. This also acts to increase contrast as electrons scattered through large angles do not contribute to the final image. However, the theoretical resolution is decreased when smaller objective apertures are used ( $\alpha$  smaller).

Image contrast in the TEM results from three mechanisms, which can all contribute to the image: mass/thickness, diffraction and phase. In brightfield imaging the objective aperture is centred around the optical axis, and the size of the aperture determines the number of scattered electrons that contribute to the image. Thicker or higher density regions will scatter electrons more and therefore appear darker in the image. If the specimen is crystalline then the electrons may be diffracted, and the contrast will depend on crystal orientation and strongly diffracting regions will appear dark. Phase contrast arises when electrons of different phase are allowed through the objective aperture and contribute to image formation. As most scattering events result in a change in phase, most images contain some phase contrast as it is impossible to select a small enough objective aperture to exclude all scattered electrons. If diffracted electron beams are allowed through this aperture and interfere, a lattice image is produced which allows interplanar spacings of crystalline materials to be directly measured from the image.

### 14.2.2 Practical aspects

The main instrumental variables in terms of the operating conditions of the TEM are the choice of accelerating voltage and objective aperture size.

Increasing accelerating voltage results in: (i) increased specimen penetration, (ii) shorter electron wavelengths (better resolution), (iii) reduced image contrast and (iv) reduced specimen damage. Increasing objective aperture size results in: (i) better theoretical resolution, (ii) reduced contrast and (iii) increased spherical aberration.

The actual conditions chosen are generally a compromise and depend on the nature of the specimen and the information required (contrast versus resolution).

TEM is a very powerful technique for obtaining information on colloidal systems. It is routinely used for obtaining information on particle size, shape, dispersity and aggregation. It can also provide analysis of internal structure, chemical composition and crystallographic information. The most critical step in the analysis of colloidal systems is often the sample preparation. As mentioned previously a number of general restrictions (vacuum, thermal and photostability) are placed on the specimen due to instrument design. For TEM analysis further limitations are imposed on specimen dimensions. The specimens are usually supported on 3 mm diameter copper mesh grids, covered with a thin film of carbon or a carbon coated polymer film. Such support films are chosen because they are of low atomic weight, and amorphous, so minimise information loss in the image. In addition the sample should generally be as thin as possible ( $<1\ \mu\text{m}$ ) to allow transmission of electrons, and to minimise beam damage.

A wide variety of techniques are available to prepare samples for analysis in the TEM (Table 14.2). However, all sample processing steps have the potential to introduce imaging artefacts. Below some selected examples of TEM studies on colloidal particles are used to highlight the range of information that can be obtained from simple dispersions of colloidal particles and the additional information that can be acquired by some of the indirect sample preparation methods.

### 14.2.3 Polymer latex particles

Suspensions of small particles can be dried directly onto grids for examination in the TEM. For particle sizing TEM is often used as a complementary technique to e.g. light scattering. Image analysis software allows a range of measurements to be performed on the particles (Figure 14.3).

**Table 14.2** Common methods of specimen preparation for TEM

---

**Small particles**

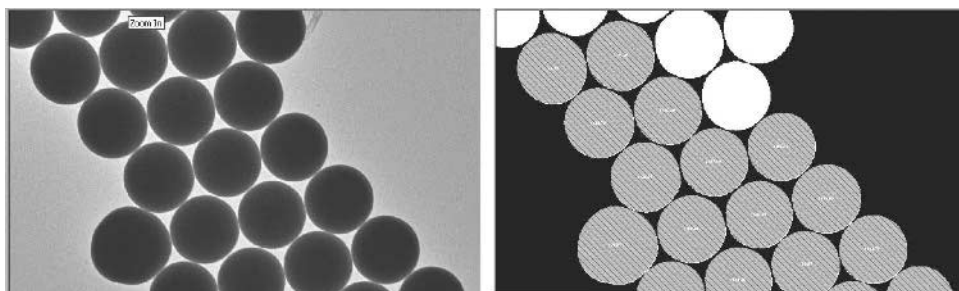
- *Evaporation (dispersions)*
- 'Dusting' (dry powders)
- Electrostatic deposition
- Freeze-etching

**Bulk materials**

- *Ultramicrotomy*
- Ion beam, chemical, electrochemical thinning

**General**

- *Staining techniques*
  - 'Decoration techniques'
  - Surface replication
-



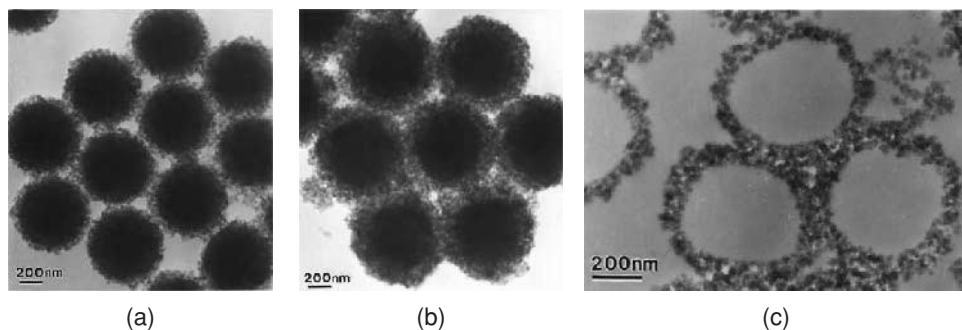
**Figure 14.3** Auto particle sizing of polymer latex particles.

Assuming the aggregation of the particles is limited (e.g. by dilution of suspension), so that individual particles can be readily distinguished, statistically significant measurements on a number of particles can be performed relatively quickly. In addition to aggregation, other possible artefacts are particle shrinkage, on drying and examination under vacuum, and beam damage. The particles shown in the figure are relatively large ( $\sim 1100$  nm) so the contrast in the projected image arises from absorption/inelastic scattering. This is fine for analysing external size and shape of these relatively stable particles. For more thermally sensitive particles this may result in ‘melting’ under exposure to the electron beam.

#### 14.2.4 Core/shell particles

TEM is particularly useful for directly imaging heterogeneous particles and allows the determination of information in addition to the external diameter. For example Figure 14.4 shows a nano-structured coating assembled onto a polymer latex particle [1].

The shell is less dense than the core and so is readily discernible in the projected image. Measurements of core diameter and shell thickness can be made, and in this particular



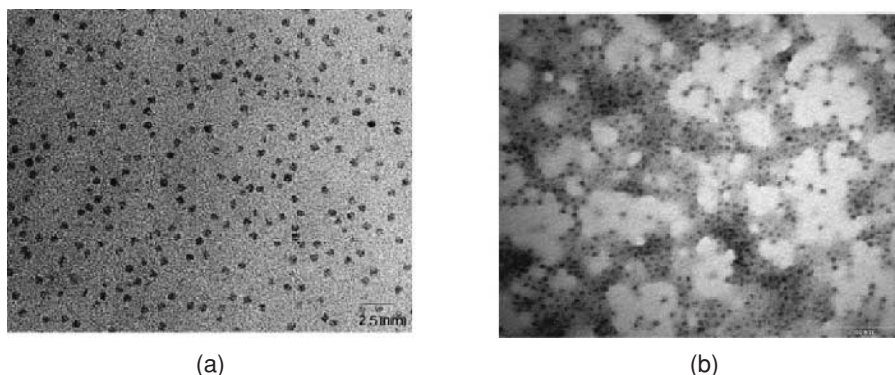
**Figure 14.4** Layer by layer assembly of zeolite particles onto latex particles (a) three layers, (b) five layers and (c) ultramicrotomed thin section of shell after core removal [1].

example used to determine the increase in shell thickness as each new layer of shell material is sequentially added. However, after removal of the core, although the shells can still be imaged, it is difficult to measure the shell thickness directly. Cross sections of the shells were prepared for TEM analysis by ultramicrotomy – a common technique for thinning bulk materials for TEM studies. The dry shell material was embedded in a polymer resin in a mould. Very thin slices (50–100 nm) of this sample were then cut using a diamond knife in an ultramicrotome. The slices were collected on a grid and images of the shells in cross section were recorded on the TEM.

Besides methods to make thin samples for TEM analysis it is often necessary to increase the contrast of a material which cannot be imaged directly. The most common method for doing this is using stains, which were developed to improve the contrast of biological material in the TEM. The basic principle is that solutions of salts of heavy metals (inherently high contrast) can be applied to TEM samples to increase contrast. Positive stains act by interacting with specific functional groups (e.g. osmium tetroxide with  $\text{C}=\text{C}$ ). Negative staining is an indirect method in which the stain solution is allowed to dry on the grid. The heavy metal salts concentrate around the low contrast material on the grid, allowing a 'negative' image to be obtained. For example Figure 14.5 shows images of the iron storage protein ferritin. Normally only the iron oxide (ferrihydrite) core is visible when examined by TEM.

If the sample is negatively stained with uranyl acetate, a halo around the dense cores can be discerned, which corresponds to the protein shell. This indirect imaging allows information on the core and shell diameter to be obtained.

However, again it must be remembered that any change in the physicochemical conditions (pH, concentration, temperature, ionic strength) during sample preparation can potentially change the nature of the colloidal system. For example, supramolecular aggregate structures of surfactants and lipids can be particularly susceptible to alteration during sample processing. For such labile systems the optimum technique is cryo-TEM, where the thin film



**Figure 14.5** TEM image of the iron storage protein ferritin: (a) image of iron oxide cores and (b) negatively stained sample to allow the protein shell to be imaged indirectly (©Mei Li, School of Chemistry, University of Bristol).

specimen is first rapidly frozen by plunge freezing in liquid ethane and then the thin vitreous ice film is imaged at low temperature (liquid nitrogen) using a cryo-TEM holder.

### 14.2.5 Internal structure

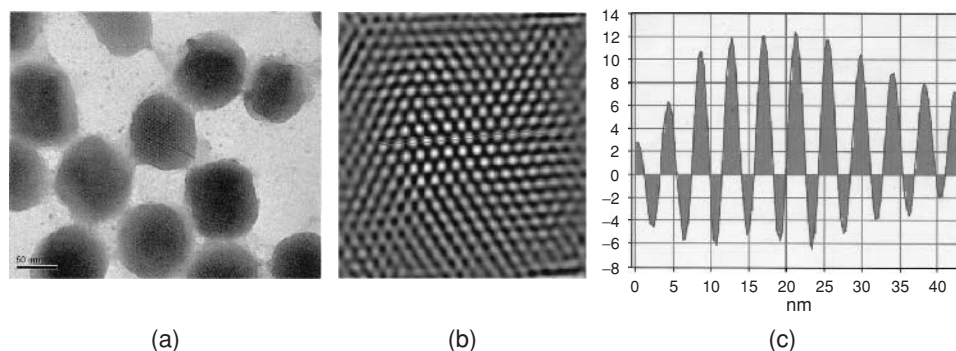
As discussed above one of the key advantages of TEM is the ability to image the internal structure of materials. TEM can be used to provide complementary information to other techniques such as surface area measurements for porous materials or XRD for crystalline materials. However the small sample volumes required for TEM means that when bulk analyses are not possible for comparison, characterisation can be provided by TEM analysis alone.

Over the last 10 years there has been considerable interest in meso-structured inorganic materials prepared using surfactant aggregate structures as templates. TEM analysis of such materials allows factors such as pore size, wall thickness, pore order, etc to be determined directly. Figure 14.6 shows a TEM image of a colloidal dispersion of silica MCM-41 particles.

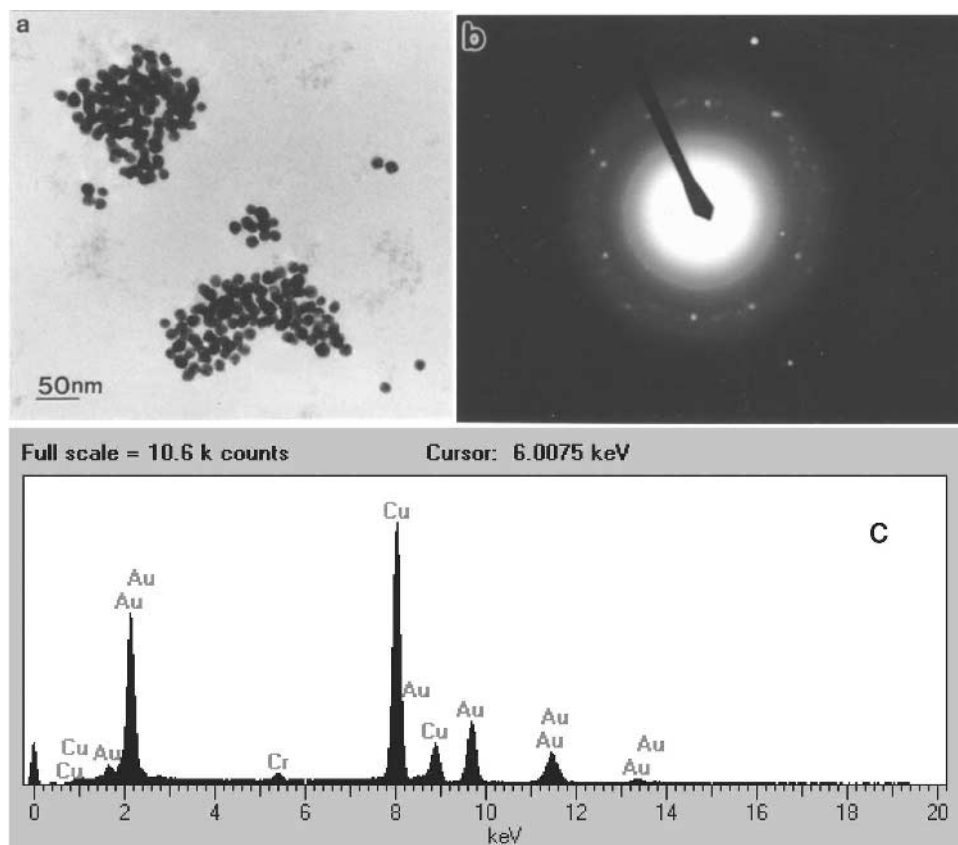
Image processing and analysis can be used to get accurate measurements of the pore–pore distance. For ceramic colloidal particles one common artefact is the sintering of aggregated particles, e.g. condensation reactions of surface silanol groups on silica particles. The related crystalline microporous zeolite materials are often extremely beam sensitive and tend to lose crystallinity on prolonged beam exposure. With these and other thermally sensitive materials beam stability is improved by just imaging at low temperature, using a cryo-TEM holder.

Crystalline colloidal materials are usually characterised by TEM analysis of a number of particles for size, shape, dispersity, composition and crystallinity (Figure 14.7).

Ultimately at high magnification the resolution of TEM allows lattice imaging of individual crystalline particles. The demand for high resolution TEM instruments has been driven in part by the explosion in interest in the synthesis and characterisation of nanostructured materials. Improvements in instrument specifications such as field emission guns, higher



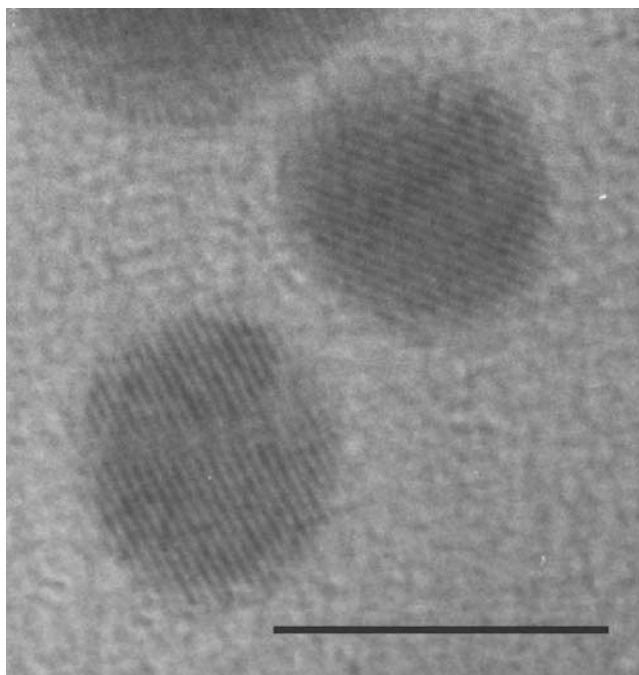
**Figure 14.6** (a) TEM image of colloidal particles of MCM-41, (b) filtered image and (c) corresponding linescan analysis of contrast variation.



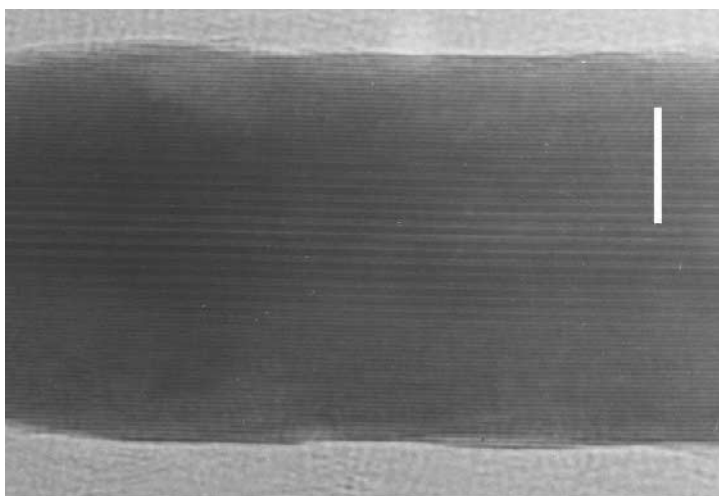
**Figure 14.7** (a) TEM characterisation of colloidal gold suspension, (b) corresponding electron diffraction pattern and (c) energy dispersive X-ray analysis spectrum (Cu peaks from support grid). Reprinted with permission from *Nature* [5], copyright 2003 Macmillan Magazines Limited.

accelerating voltages, aberration corrected lenses and higher resolution digital cameras have led to HRTEM becoming a more routine characterisation technique.

Figure 14.8 shows a high resolution image of gold nanoparticles ( $\sim 3$  nm). The fringes in the image correspond to the  $\{200\}$  lattice planes of gold (spacing = 0.2039 nm). In addition to the high resolution imaging of individual particles it is also routinely possible to obtain diffraction and X-ray analysis from isolated nanoparticles using electron beams of nominal diameter  $< 10$  nm. This is particularly useful when characterising particles of varying size, shape or composition. For example Figure 14.9 shows a HRTEM image of a gold rod prepared from the particles shown in Figure 14.8 [2] by a seeded growth technique. From such images it was possible to determine that the rods were not single crystalline particles, but were in fact multiply twinned. Ultimately such insights allow for the improvement and optimisation of the synthesis to improve the yield of the desired anisotropic rod-like particles.



**Figure 14.8** HRTEM image of gold particles (scale bar = 5 nm).



**Figure 14.9** HRTEM image of gold nanorod viewed down the  $\langle 112 \rangle / \langle 100 \rangle$  zone showing continuous  $\{111\}$  fringes ( $d = 0.236$  nm) parallel to the direction of elongation. The fringes are modulated in the central region of the twinned crystal into wider stripes due to double diffraction arising from the superposition of twin domains aligned along different zones (scale bar, 5 nm [2]).



## 14.3 Conventional SEM

### 14.3.1 Background

In SEM the electron beam is scanned across the specimen surface point by point. The signal collected from each point is used to construct an image on the display, with the cathode ray tube beam and the column beam following a synchronised scanning pattern. Thus the displayed image is the variation in detected signal intensity as the column beam is scanned across the sample. The ultimate performance of the SEM is limited by the beam diameter. The lenses in the SEM do not magnify the image, they demagnify the beam. The condenser lens reduces the beam diameter from 50  $\mu\text{m}$  to  $\sim 5$  nm. The image is focused by adjusting the final lens such that the beam has the minimum diameter at the specimen surface. The magnification is given by the simple relationship between the area of specimen scanned relative to the area of displayed image.

### 14.3.2 Types of signal

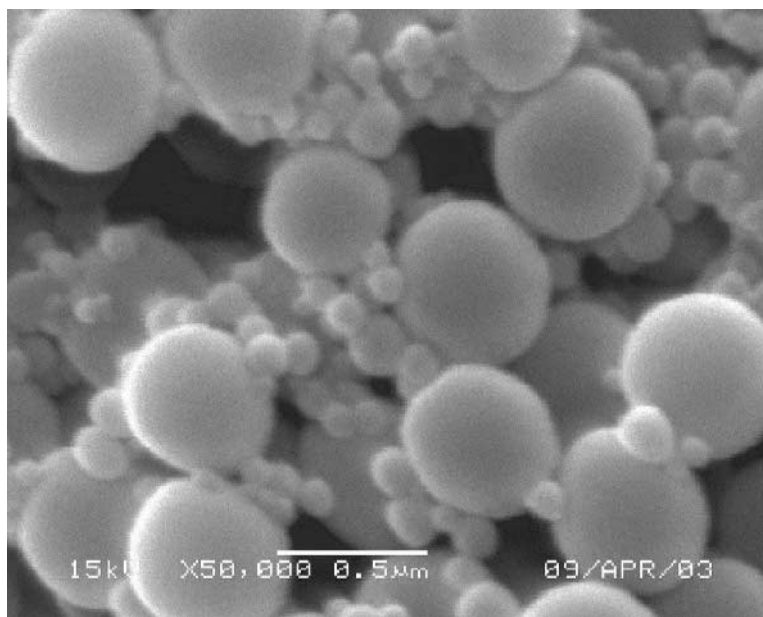
Secondary electrons are produced when incident beam electrons knock out loosely bound conduction electrons. Due to the low energy of secondary electrons ( $< 50$  eV) they can only escape if they are within  $\sim 10$  nm of surface (Figure 14.2). The detected signal intensity depends on the angle between the beam and the specimen. These two factors mean that the secondary electron signal provides the highest resolution topographic information.

In contrast the backscattered signal is produced by elastically scattered electrons, deflected through angles between  $0^\circ$  and  $180^\circ$  by atoms within the specimen. Those scattered through angles greater than  $90^\circ$  can re-emerge from the specimen surface still with a high energy. Under similar operating conditions the signal will be produced from a larger volume than the secondary electron signal and so will give lower resolution topographical information. However, scattering events are more likely with atoms of higher atomic weight (or if the incident electron has low energy), so the signal can be used to give qualitative compositional information in heterogeneous samples.

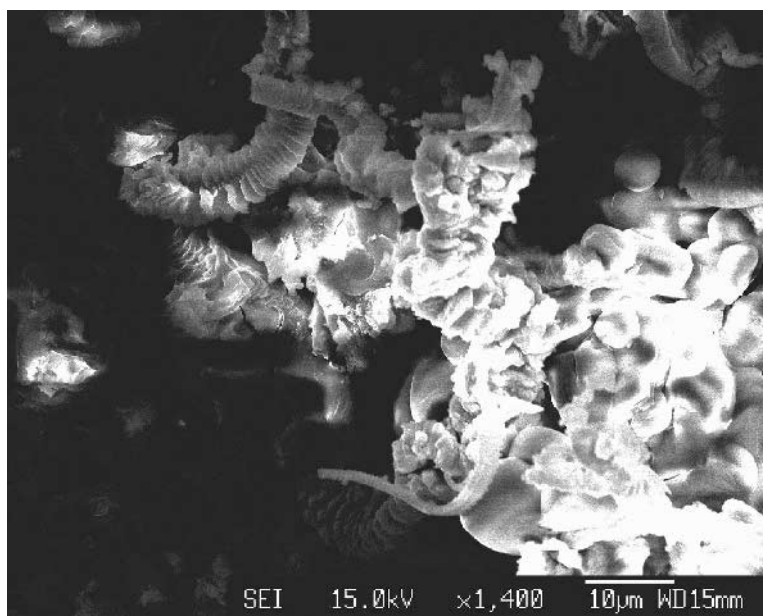
### 14.3.3 Practical aspects

SEM is used to look at the surface structure of materials with a resolution  $< 2$  nm. Thus it can be used for determining particle size, shape and dispersity (Figure 14.10) as well as chemical composition and distribution from the characteristic X-ray signal.

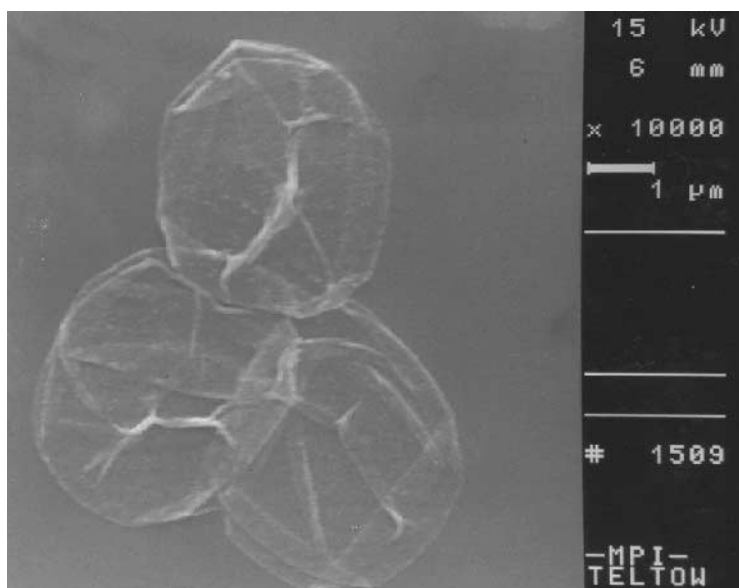
In the SEM, unlike the TEM, the sample dimensions are restricted purely by the physical size of the column rather than the sizes of lenses and apertures. Typically sample holders are in the region of 10–40 mm in diameter, and obviously the specimens can be much thicker. It is generally as easy to look at bulk specimens as it is at thin films or dispersions. Besides, the general stability criteria of the specimens that can be routinely imaged, non-conducting specimens have to be coated in a thin conductive film (C, Au, Pt/Pd). This prevents charge build-up on the specimen and the associated image distortion (Figure 14.11).



**Figure 14.10** SEM image of polydisperse latex sample.



**Figure 14.11** Powdered sample showing effects of specimen charging on image quality.



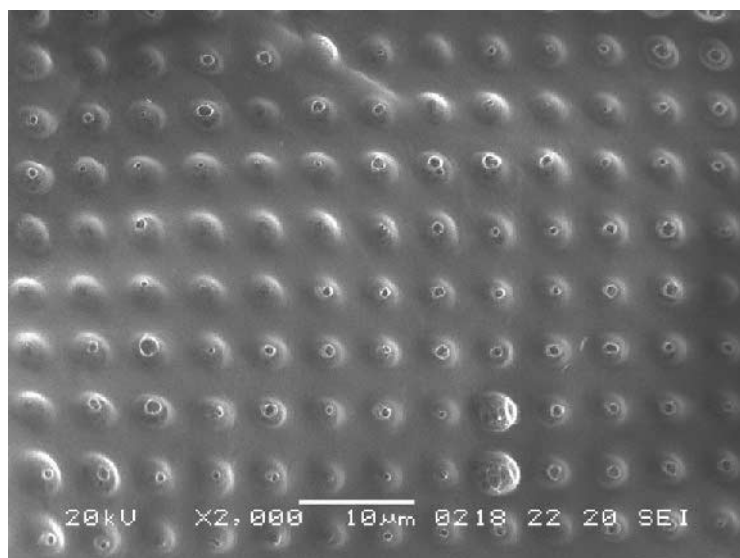
**Figure 14.12** Thin polymer shells appear 'transparent' when imaged at high accelerating voltage [3].

The main instrumental operating conditions that can be varied are accelerating voltage, scan speed, spot size/beam current, aperture size, working distance and tilt.

The accelerating voltages used in the SEM are lower than those for TEM, typically between 1 and 30 kV. The choice of voltage depends on the nature of the specimen material and the magnification range and image resolution required. Generally for high resolution work small beam diameters are needed, thus high accelerating voltages are used to produce a good emitted signal for image formation. However, high accelerating voltages in SEM result in increased specimen penetration (e.g. Al 5 keV – 1 nm, 30 keV – 10 nm). Image information is produced from deeper within the specimen, so surface detail is lost and the chance of specimen damage is increased (Figure 14.12) [3].

One of the main advantages of SEM over light microscopy is the increased depth of field (range of positions of object for which eye can detect no change in sharpness of image). For example, at an image magnification of  $100\times$ , the depth of field in the SEM is  $\sim 1$  mm, compared with  $1\ \mu\text{m}$  in the optical microscope. A small aperture gives larger depth of field, as well as increased resolution. Short working distances are also used for high resolution work, but this reduces the depth of field. For backscattered electron imaging and X-ray analysis larger apertures and beam diameters are used to increase the beam current and hence yield of signal. Again, the increased beam current can lead to specimen damage (Figure 14.13) [4].

The scan speed is essentially decreased to improve signal to noise ratio when recording images and focusing. For finding areas of interest a fast refresh rate is used (25 frames/s). The yield of secondary electrons detected can be improved by tilting the specimen (Figure 14.14), but this causes image distortion (however, particle sizes can still be determined from the images).

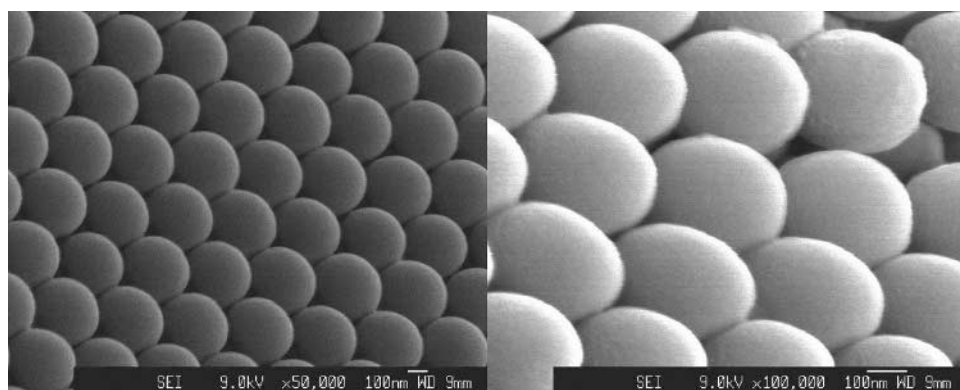


**Figure 14.13** Damage to surface of a starch gel induced by high beam currents [4].

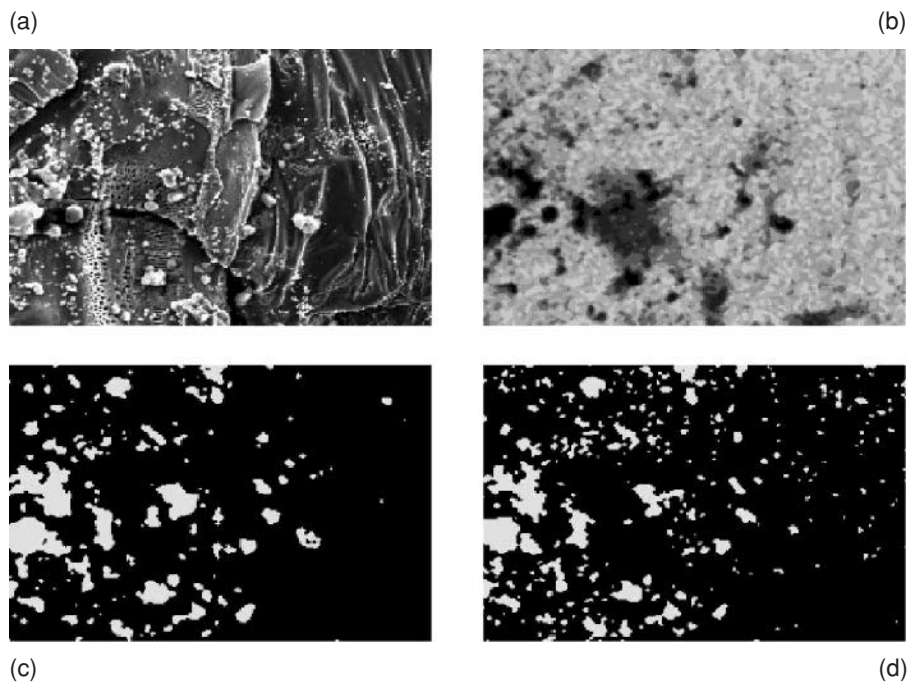
Another advantage of the SEM is that the characteristic X-rays detected can be plotted as an elemental composition image (Figure 14.15). Due to the larger volume from which the X-rays emerge from within the sample the resolution of these elemental maps is less than the corresponding SEM image. However, they can prove useful for determining the homogeneity of samples.

In the SEM the use of field emission guns has resulted in improved resolution (Figure 14.16). With a resolving power of the order of a few nm, it is possible to routinely image individual particles within aggregate structures.

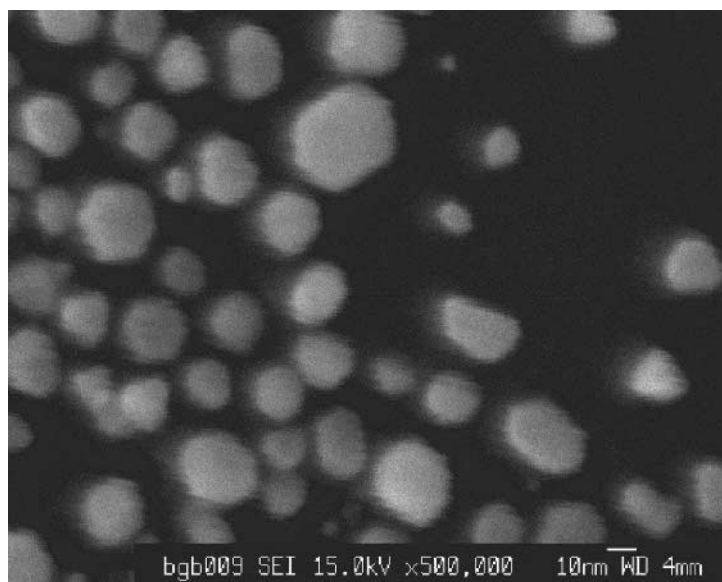
For example, the homogeneity of coverage of nanostructured thin films on solid substrates can be ascertained without all the associated sample preparation required for TEM analysis.



**Figure 14.14** Tilting spherical polymer latex specimen to improve signal results in image distortion.



**Figure 14.15** Calcium phosphate precipitates grown on a titania substrate: (a) original image, (b) titanium, (c) calcium and (d) phosphorus images.

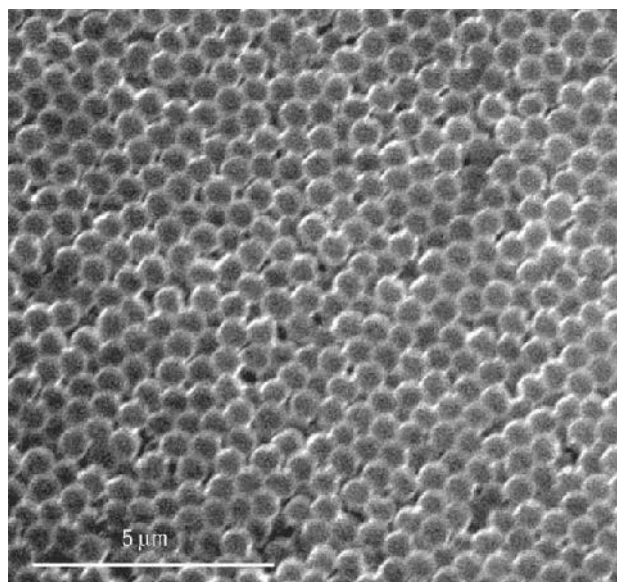


**Figure 14.16** FEG-SEM image of a resolution test specimen (gold on carbon).

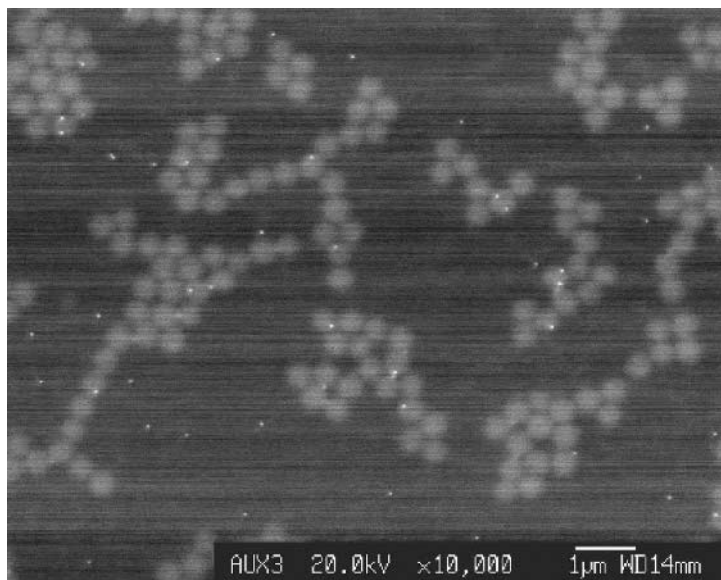
Over the last decade a number of improvements to instrument design have been associated with developing non-conventional instruments which can image samples in their native state. Traditional methods of preparing hydrated samples for EM analysis include freeze-drying and critical point drying. Although these techniques reduce structural damage caused by surface tension effects associated with air-drying wet samples, both have limitations. For example, the solvent exchange steps required to dehydrate samples prior to critical point drying (and also embedding material for sectioning) can result in some loss of structure or solubilisation of certain components. The technique of cryo-SEM can minimise further the production of artefacts or loss of structure during specimen processing. Like cryo-TEM the first step is rapid freezing of the sample. The specimen is then transferred into a cooled specimen chamber, sputter coated and imaged.

The development of environmental SEMs has expanded further the range of information that can be obtained from samples [5]. The gun is still maintained at high vacuum but a differential pumping system allows a low pressure of gas around the sample. Secondary electrons are used for image formation, and the resolution of these instruments is  $\sim 5$  nm. The advantage of environmental chambers is that for insulators no conductive coating is required. In addition, if the gas is water vapour hydrated specimens can be imaged. Controlling the temperature allows the state of hydration to be varied, thus dynamic processes such as aggregation and film formation can be studied (Figure 14.17).

Finally a novel specimen holder has recently been developed to allow the imaging of samples in their hydrated state [6]. The specimen capsules have a vacuum resistant, electron transparent membrane which also permits compositional analysis by EDX. Small volumes of suspension are applied to the membrane, and then the holder is sealed. Images are produced using high energy backscattered electrons (secondary electrons produced from the



**Figure 14.17** ESEM image of partially dehydrated film forming latex particles prior to coalescence (A.M. Donald, (2003) *Nature Mater.*, **2**, 511, reproduced with permission).



**Figure 14.18** BSE images of an aqueous suspension of 200 nm silica and 30 nm gold particles (Quantomix capsule).

specimen are absorbed by the membrane). The detected signal is produced from particles close to or adsorbed on the membrane. Potential applications include imaging cells, emulsions, suspensions, creams, etc., with the obvious limitations of solvent compatibility and adherence to the membrane. An example is shown in Figure 14.18.

## 14.4 Summary

Although conventional SEM and TEM are well established techniques, improvements in microscope design and new techniques continue. Currently several of these developments are being driven by the differing EM requirements of bio and nanotechnologists. However, colloid scientists stand to benefit from these improvements as the range of techniques available for imaging soft matter and small particles increases. The following texts are useful general references.

1. Goodhew, P.J. (2001) *Electron Microscopy and Analysis*. 3rd edition Taylor and Francis, London.
2. Chescoe, D. and Goodhew, P.J. (1990) The operation of transmission and scanning electron microscopes. *Royal Microscopical Society Handbook*, vol. 20. Oxford University Press, Oxford.
3. Environmental Scanning Electron Microscopy, Philips.
4. The Principles and Practice of X-ray Microanalysis, Oxford Instruments.
5. A Guide to Scanning Microscope Observation, Jeol.  
[www.matter.org.uk](http://www.matter.org.uk)

## References

1. Davis, S.A., Breulmann, M., Rhodes, K.H., Zhang, B. and Mann, S. (2001) *Chem. Mater.*, **13**, 3218–3226.
2. Johnson, C.J., Dujardin, E., Davis, S.A., Murphy, C.J. and Mann, S. (2002) *J. Mater. Chem.*, **12**, 1765–1770.
3. Donath, E., Sukhorukov, G.B., Caruso, F., Davis, S.A. and Moehwald, H. (1998) *Angewandte Chemie*, Int. Ed., **37**, 2202–2205.
4. Zhang, B., Davis, S.A. and Mann, S. unpublished results
5. Donald, A.M., (2003) *Nature Mater.*, **2**, 511.
6. [www.quantomix.com](http://www.quantomix.com)



# Index

- $\chi$  parameter, 110
- $\zeta$ -potential, 15, 42, 44
  
- absorption cross sections, 237
- acacia, 224
- accumulation mode, 181, 182
- achromatic (spherical) aberrations, 268
- acid rain, 180, 184
- additives, 221
- adhesion, 165
- adhesive failures, 166
- adsorbed amount, 119
- adsorption, 116
- adsorption isotherm, 131, 146
- aerodynamic
  - diameter, 186
  - lens, 187, 188
  - particle sizer, 190
  - size, 190
- aerosol, 2, 180
- affine, 201
- aggregates, 153
- aggregation, 156
- Aitken nuclei, 181
- amphiphilic, 50
- anisotropic particles, 244
- antifoamers, 168
- anti-thixotropic, 209
- applied potential, 17
- area per molecule, 252
- Asakura–Oosawa, 151
- association colloids, 1, 8
- atomic force microscopy, 145
- atomistic
  - modelling, 118
  - simulation, 113
- attractive forces, 225
- average particle separation, 5
- axisymmetric drop shape analysis, 176
  
- backscattered electrons, 266
- Bancroft's rule, 81, 168
  
- Bashforth and Adams, 177
- BET, 116
- Bingham, 211
- blade coating, 204, 210
- blob, 123
- block copolymer, 136, 223
- Boltzmann law, 263
- bound fraction, 119, 131
- Bragg peaks, 234
- bridging
  - flocculation, 156
  - interactions, 144, 155
- Brownian
  - collision, 42
  - forces, 262
  - motion, 225
- brush, 140
  
- calcium phosphate, 279
- captive bubbles, 176
- carboxylic acid, 223
- cascade impactor, 187, 188, 189, 190
- Casson, 211
- cavitation, 7
- central region, 128
- characteristic time, 207
- charge stabilisation, 143
- charge stabilised suspensions, 15
- charging mechanism, non-polar media, 6
- chromatic aberrations, 268
- clays, 16
- cloud
  - condensation nuclei counters, 189
  - point, 63
- coacervate, 224
- coagulation, 45, 183
- coalescence, 11
- coarse mode, 181, 195
- cohesion, 165
- cohesive failures, 166
- colloidal
  - crystals, 12
  - dispersion, 1

- colloidal (*Continued*)
  - gels, 263
  - interactions, 219
- comminution, 6
- compliance, 214
- concentrated dispersions, 11, 245
- condensation, 99
- condensation particle counter, 189
- cone and plate, 206
- configurational entropy, 117
- contact angle, 161, 166, 176
- continuous phase, 1
- contour length, 102
- contrast, 239
- contrast variation, 246
- copolymerisation, 223
- copolymers, 100, 111, 136
- cream, 213
- creep, 210
- creep compliance, 214, 215, 216
- critical
  - coagulation concentration, 43
  - micelle concentration, 55, 63
  - molecular weight, 220
  - packing parameter, 69, 70
  - stress, 225
  - surface tension, 172
- Cross, 211
- cubic phase, 73
- cup and bob, 206
- cyclones, 187, 188
- de Gennes, 105, 122, 149
- Deborah number, 212
- Debye, 39
- Debye length, 24, 38, 42
- Debye–Hückel approximation, 24
- degree of polydispersity, 107
- deliquescence point, 182
- depletion interactions, 144, 151
- deposition, 183
- Derjaguin approximation, 145
- detector, 236
- detergency, 167, 168
- deuterium moderator, 235
- differential
  - capacitance, 17
  - light scattering, 194
  - mobility analyser, 192
  - mobility particle sizer, 192
- diffraction limited, 266
- diffuse layer, 22
- diffusion
  - batteries, 191
  - coefficient, 9, 233
  - constant, 46, 47
  - denuder, 195, 196
- dilute dispersion, 239
- DIN, 206
- disc, 226
- dispersant, 15
- dispersed phase, 1
- dispersion
  - dry powder, 186
  - forces, 172
- distal region, 128
- DLVO, 10, 40
- DNA, 106
- Doppler, 233
- drag, 205
- Du Nouy, 176
- Dupré equation, 165
- dynamic surface tension, 56
- EC, 195–7
- EDX, 280
- effective particle volume fraction, 11
- effectiveness of adsorption, 61
- efflorescence point, 182
- elasticity, 212
- electrical
  - double layer, 7
  - mobility size, 190, 191, 192
- electroacoustics, 35
- electrocapillary, 19
- electrokinetics, 29
- electron
  - densities, 230, 245
  - microscopy, 228, 266
- electro-osmosis, 31
- electrophoresis, 32
- electrophoretic
  - mobility, 10
  - velocity, 10
- electrostatic levitation, 198
- elemental carbon (EC) fraction, 182
- emulsion, 2, 139, 205, 213
- end to end distance, 108
- end-segment weighting factors, 125
- energy
  - dissipation, 213
  - stored, 213
- entanglement, 100, 209, 220, 223
- enthalpy of micellisation, 67
- environmental SEMs, 280
- epiphaniometer, 190
- equilibrium concentration, 146
- evaporation, 206
- exact enumeration, 119
- excluded volume, 104, 105, 115
- factors affecting CMC, 67
- ferritin, 271
- field emission, 278
- film bending rigidity, 88
- filters, 188, 189, 190, 195, 196
- fine particles, 180
- flocculate, 113, 226
- flocculation, 45
- Flory, 102, 108

- Flory surface parameter, 117  
Flory–Huggins  
  lattice, 109  
  parameter, 109, 115, 147  
flow  
  cup, 203, 204  
  curves, 211  
foam, 2, 205  
forces normal, 206  
form factor, 240  
fractal surface, 249  
fractals, 249  
fracture, 206  
free  
  energy, 149  
  radical, 99  
frequency, 212  
frictional coefficient, 9  
FTIR, 133  
functionalisation, 221  
  
Galvani potential, 16  
Gaussian curvature, 86  
gel, 153, 226  
gelatine, 224  
geometric mean theorem, 173  
geometries, 204, 205, 206  
Gibbs  
  dividing surface, 56  
  equation, 58  
  pre-factor, 59  
gold, 279, 281  
gold nanorod, 274  
good solvent, 115  
Gouy–Chapman theory, 22  
gradient force, 257  
gravitational force, 217  
Guinier  
  law, 241, 245  
  plot, 241  
  
Hamaker constants, 38, 163, 172  
hanging  
  bubbles, 176  
  drops, 176  
hectorite, 226  
Helmholtz model, 19  
Herchel–Bulkley, 211  
hexagonal phase, 72  
high shear rate viscosity, 211  
holographic optical tweezers, 260  
homologous, 172  
Hookean, 202  
HRTEM, 273  
Huggins, 108  
hydrodynamic, 225  
  coupling, 12  
  layer, 133  
  radius, 233  
hydrophilic, 5  
  hydrophilic–lipophilic balance (HLB), 84  
  hydrophobic, 6, 51, 224, 225  
  hydroxycellulose, 224  
  
ice cream, 1  
ideally polarisable, 17  
ILL, 237  
image contrast, 268  
impaction, 183, 185, 187, 189  
impactor, 187, 190, 195, 196  
inert ions, 16  
inertia, 214  
initiation, 98, 99  
inlet, 186, 187, 188  
inner Helmholtz plane, 28  
instantaneous compliance, 214  
inter-particle  
  forces, 10  
  interference, 239  
  potential, 113  
interfacial  
  area, 5  
  charge, 6  
  region, 5  
  structure, 5  
  tension, 78, 176  
  thickness, 6  
intermolecular forces, 172  
internal structure, 272  
intrinsic viscosity, 108  
ion chromatography, 195, 196  
ionic  
  atmospheres, 39  
  methods, 99  
  surfactants, 15  
ionisation of surface groups, 15  
ISIS, 235  
isoelectric point, 15  
isokinetic sampling, 186  
isomorphic substitution, 6, 16  
isotopes, 230  
  
Kelvin effect, 182, 194  
kinetic stability, 42  
Knudsen number, 183  
Köhler theory, 182  
Krafft temperature, 62  
Kreiger, 211  
Kuhn, 104  
Kuhn lengths, 104  
  
lamellar phase, 72  
Langmuir, 116, 128  
Laplace equation, 176  
Laponite, 226  
laser Doppler electrophoresis (LDE), 32  
laser trap, 258  
latex, 225, 276  
latex particles, 280  
layer thickness, 133, 251

- levelling, 216
- light microscope, 266
- light scattering, 228, 231
- linearity, 202, 212, 213
- lines of flow, 207
- Lippman equation, 18
- liquid, 212
- liquid crystalline mesophases, 70
- liquid/liquid interface, 139, 165
- log jamming, 211
- London, 37
- loops, 113
- Lorentzian decay, 263
- loss modulus  $G''$ , 213
- low shear rate viscosity, 211
- lubricating layer, 218
- marginal solvents, 149
- Mark–Houwink parameters, 108
- mass concentration, 188, 189
- maximum bubble pressure, 176
- maximum packing fraction, 220, 221
- mean curvature, 86
- mean square
  - displacement, 263
  - end to end distance, 103
- measuring geometries, 204
- mercury electrode, 17
- metal oxides, 15
- metal/electrolyte interfaces, 17
- micelles, 9, 224
- microemulsion phase behaviour, 90
- Mie theory, 233
- model, 151
- molecular
  - dynamics, 121
  - liquid, 208
- momentum, 255
- monochromatic, 236
- monodisperse polymer, 107
- monolayer, 251
- Monte Carlo, 118, 121
- morphology dependent resonances, 199
- multi-body interactions, 11
- mushroom, 121, 140
- microemulsions, 77
- nanofabrication, 260
- nanotechnology, 2
- nebuliser, 185, 186
- nephelometer, 194
- networks, 221, 224
- neutron
  - moderator, 236
  - reflection, 250
- Newton's dot, 203
- Newtonian, 203, 207
- NIST, 237
- NMR, 132
- non-equilibrium states, 153
- non-wetting, 162
- nucleation
  - and growth, 8
  - particles, 181
- number
  - average molecular weight, 106
  - concentration, 188, 189
  - density, 4
- OC, 195, 196
- optical
  - forces, 256
  - levitation, 198
  - particle counter, 193, 194
  - size, 190, 193
  - traps, 257
  - tweezers, 257
- organic carbon (OC) fraction, 182
- oscillation, 212
- osmium tetroxide, 271
- osmotic pressure, 111, 114
- Ostwald ripening, 11
- outer Helmholtz plane, 25
- Owens–Wendt, 174
- packing parameter, 84
- pair potential, 36
- PALS, 33
- pancake, 121, 140
- partial wetting, 162
- particle, 220
  - additives, 225
  - aggregation, 205
  - jamming, 205
  - shape, 242
  - size, 240, 269
  - size distribution, 242
- paste, 218
- PCS, 133
- Peclet number, 225
- perfect wetting, 162
- periodic boundary, 122
- persistence, 104
- persistence length, 106
- phase, 213
  - diagram, 226
  - modulator, 260
  - rule, 91
  - transitions, 153
- phase analysis light scattering (PALS), 32
- Phase Inversion Temperature (PIT), 84
- photoacoustic soot sensor, 197
- photon correlation spectroscopy, 233
- photonic band-gap materials, 260
- Pickering emulsions, 167
- plastic, 211, 217
- plastic models, 211
- pluronic, 139
- PMMA, 153
- point of zero charge, 19

- Poisson equation, 23  
Poisson–Boltzmann equation, 23  
polarization factor, 231  
poly(acrylic acid), 106, 156  
poly(ethylene oxide), 128, 156  
poly(hydroxystearic acid), 149  
poly(methylmethacrylate), 149  
poly(propylene oxide), 137  
poly(vinyl chloride), 99  
poly(vinyl pyrrolidone), 132  
polyacrylate, 219  
polydimethylsiloxane, 101  
polydispersity, 3, 99, 116, 131, 219, 242  
polyelectrolytes, 106, 155  
polyethylene, 103, 163, 174  
polyethylene oxide, 224  
polymer, 220, 223  
    adsorption, 146  
    latex, 278  
    networks, 224  
    viscosity, 107  
polymerisation, 98  
polymethylmethacrylate, 174  
polystyrene, 45, 101, 118, 119, 128, 141, 156  
polytetrafluoroethylene, 163  
poor solvent, 105  
Porod's law, 247, 249  
potential-determining ions, 16  
pour point, 203  
power law, 211  
primary  
    aerosols, 182  
    maximum, 42  
propagation, 98  
proton induced X-ray emission, 195  
proximal region, 126  
pseudo-binary phase diagram, 95  
pseudoplastic, 211, 215  
PTFE and polystyrene, 173  
pulsed neutron source, 237  
  
quadrant photosensor, 261  
  
radius of gyration, 103, 104, 232  
Raman scattering, 199, 200  
random copolymer, 112, 136  
random walk, 102  
rate of coagulation, 46  
Rayleigh, 231  
Rayleigh–Gans–Debye theory, 233  
reflectivity, 251  
reflectometer, 251  
relative  
    deformation, 202  
    humidity, 182  
relaxation time, 212  
resolution, 268  
respiratory tract, 180  
retardation time, 215  
reversible aggregation, 45  
  
Reynolds number, 207  
RMS thickness, 135  
Rod, 226  
Russel, 149  
Rutherford Appleton Laboratory, 235  
  
SANS, 128, 234  
SAXS, 234  
scaling theory, 126  
scanning mobility particle sizer, 193  
scattering, 228  
    angle, 229  
    force, 256  
    length, 237  
    length density, 237, 239  
    vector, 229  
Scheutjens and Fleer theory, 123  
secondary  
    aerosols, 182  
    electrons, 266, 280  
    flow, 207  
sedimentation, 143, 216  
sedimentation velocity, 10  
self-avoiding walk, 120  
self-similar, 126  
SEM, 266  
sessile drops, 176  
shear  
    history, 210  
    modulus, 202  
    plane, 29  
    rate, 203  
    thinning, 209  
    thickening, 210  
    viscosity, 10, 203  
Shultz–Hardy rule, 43, 44  
silica, 210, 272, 281  
silver halide, 16  
single particle time-of-flight mass  
    spectrometer, 197  
size distribution, 188, 190, 191, 193, 194  
slurries, 206  
small angle scattering, 234  
small angle X-ray, 228  
soft matter, 256  
sol, 2  
solid, 212  
solvent NMR relaxation, 133  
solvent–polymer interaction energy, 108  
space filling, 226  
span, 119  
sparingly soluble salts, 16  
specific viscosity, 107  
specifically adsorbed ions, 28  
spinning drop, 176  
spontaneous curvature, 86  
spontaneous spreading coefficient, 164  
spraying, 204  
spreading coefficient, 170  
stability, 218, 225

- stability ratio, 48
- staining techniques, 269
- standard deviation, 4
- standard free energy of micellisation, 66
- starch gel, 278
- steady state, 207, 209
- steric
  - interaction energy, 147
  - interactions, 150
  - stabilisation, 144, 145, 148
  - stabiliser, 148
- Stern
  - modification, 25
  - potential, 44
- Stokes diameter, 190
- Stokes–Einstein, 233
- storage
  - modulus, 213
  - stability, 216, 217
- strain, 202, 214
- strain rate, 203
- streaming
  - current, 30
  - potential, 30
- stress, 202, 213
  - overshoot, 209
  - sweep, 213
- strong aggregation, 12
- structural forces, 10
- structure factor, 245
- styrene, 99
- supramolecular solutions, 1
- surface
  - $\chi$  parameter, 145
  - active agents, 50
  - activity, 55
  - charge, 14
  - energy, 159, 169
  - excess, 56
  - forces, 144
  - free energy, 161
  - potential, 17, 39
  - pressure, 55, 173
  - structure, 250
  - tension, 17, 55, 159, 160, 164, 169, 174, 176
- surfactant, 219, 224
  - adsorption, 169
  - adsorption efficiency, 60
  - micelles, 155
- sweep, 213
- synchrotron radiation, 235
- syneresis, 218
- tails, 113, 133
- Taylor number, 207
- TEM, 266
- terminally attached chains, 118
- termination, 98
- ternary phase diagrams, 92
- thermal energy, 208
- thermodynamic
  - of micellisation, 63
  - stability, 12
- thermotropic, 71
- thixotropic loops, 210
- thixotropy, 209
- three-phase line, 161
- titania, 15
- total internal reflection microscopy, 145
- tractor beams, 255
- trains, 113
- transverse force, 257
- ultracentrifuge, 228
- ultramicrotomy, 270
- ultrasonic radiation, 7
- van der Waals, 10, 37, 224
- van der Waals attractions, 144
- van't Hoff equation, 111
- variance, 4
- velocity selector, 236
- vibrating orifice aerosol generator, 186
- virtual impactor, 188
- viscoelastic
  - fluid, 215
  - solid, 216
- viscoelasticity, 212
- viscometers, 210
- viscosity, 228
- Volta potential, 17
- volume fraction, 4, 11, 125
- vorticity, 207
- wall slip, 205
- weak
  - attractions, 226
  - reversible aggregation, 12
- weight
  - average, 107
  - fraction, 4
- wettability envelopes, 174
- Winsor, 81
  - classification, 80
  - R-ratio, 82
- work
  - of adhesion, 165
  - of cohesion, 165, 170
- worm-like chains, 104
- X-ray
  - analysis, 273
  - fluorescence, 195
- yield stress, 211, 216, 217, 225
- Young's equation, 161, 164, 170
- zero shear viscosity, 216
- zeta potential, 29
- Zisman plot, 172